Journal of Healthcare Engineering

Artificial Intelligence in E-Healthcare and M-Healthcare 2021

Lead Guest Editor: Xingwang Li
Guest Editors: Han Wang, Wali Ullah Khan, Basem M. Elhalawany, and Thomas Aaron Gulliver

Artificial Intelligence in E-Healthcare and M-Healthcare 2021

Journal of Healthcare Engineering

Artificial Intelligence in E-Healthcare and M-Healthcare 2021

Lead Guest Editor: Xingwang Li
Guest Editors: Han Wang, Wali Ullah Khan, Basem M. Elhalawany, and Thomas Aaron Gulliver

Contents

CXR-RefineDet: Single-Shot Refinement Neural Network for Chest X-Ray Radiograph Based on Multiple Lesions Detection
Cong Lin, Yongbin Zheng, Xiuchun Xiao, and Jialun Lin
Research Article (11 pages), Article ID 4182191, Volume 2022 (2022)

Solving an Infectious Disease Model considering Its Anatomical Variables with Stochastic Numerical Procedures
Zulqurnain Sabir, Muhammad Asif Zahoor Raja, and Yolanda Guerrero Sánchez
Research Article (12 pages), Article ID 3774123, Volume 2022 (2022)

Stroke Disease Detection and Prediction Using Robust Learning Approaches
Tahia Tazin, Md Nur Alam, Nahian Nakiba Dola, Mohammad Sajibul Bari, Sami Bourouis, and Mohammad Monirujjaman Khan
Research Article (12 pages), Article ID 7633381, Volume 2021 (2021)

Hindawi
Journal of Healthcare Engineering
Volume 2022, Article ID 8926335, 7 pages
https://doi.org/10.1155/2022/8926335

Research Article

Active Management of Labor Process under Smart Medical Model Improves Vaginal Delivery Outcomes of Pregnant Women with Preeclampsia

Siming Xin,[1] **Xianxian Liu**,[2] **Jiusheng Zheng**,[1] **Hua Lai**,[1] **Jiao Zhou**,[1] **Feng Zhang**,[1] **Xiaoying Wu**,[1] **Ting Shen**,[1] **Lin Xu**,[3] **and Xiaoming Zeng**[1]

[1]*Department of Obstetrics, Jiangxi Provincial Maternal and Child Health Hospital, Nanchang, China*
[2]*Key Laboratory of Women's Reproductive Health of Jiangxi Province, Jiangxi Provincial Maternal and Child Health Hospital, Nanchang, China*
[3]*Department of Science and Education, Jiangxi Provincial Maternal and Child Health Hospital, Nanchang, China*

Correspondence should be addressed to Ting Shen; sintyst508@126.com, Lin Xu; xulinlin5270@163.com, and Xiaoming Zeng; 18070038675@163.com

Received 12 February 2022; Revised 9 March 2022; Accepted 26 March 2022; Published 7 April 2022

Academic Editor: Han Wang

Background. In a global environment of increasing cesarean delivery rate, promoting vaginal delivery, reducing the rate of first cesarean section, and the incidence of vaginal delivery complications are the objectives of obstetric medical quality and safety in China. As a common obstetric complication, preeclampsia affects the safety of many pregnant women. It is the obstetrician's great responsibility to promote vaginal delivery and improve delivery outcomes in preeclampsia. To this end, we explored the roles of active labor management under the smart medical model in improving the outcomes of vaginal delivery for pregnant women with preeclampsia. *Methods*. The clinical data of 219 cases of preeclampsia pregnant women who delivered vaginally in our hospital from January 2017 to December 2020 were retrospectively analyzed. According to different labor process management, they were divided into study group (active labor process management group) and control group (normal labor process management group). Active labor process management methods included intrapartum ultrasound, central fetal heart rate monitoring, Doula delivery, labor analgesia, and quality of life care. The differences in delivery process, delivery outcome, bleeding causes, and hemostatic measures were compared between the two groups. *Results*. (1) The incidence of preeclampsia in our hospital showed an increasing trend in recent four years; (2) in smart hospitals, the active management of labor process reduced the probability of transferring to the cesarean section in preeclampsia pregnant women with vaginal trial failure; and (3) active labor process management reduced the rate of lateral episiotomy, decreased the postpartum hemorrhage volume within two hours, and improved the vaginal delivery outcome of preeclampsia pregnant women. *Conclusions*. In the era of the rapid development of the Internet, vigorously promoting the construction of smart hospitals and actively managing the delivery process can reduce the failure rate of vaginal trial delivery and improve the outcomes of vaginal delivery in preeclampsia women.

1. Introduction

In recent years, with the rapid development of Internet technology, emerging technologies such as artificial intelligence, informatization, big data, and cloud computing have emerged, driving technological changes in the medical field. Smart medicine is an innovative application of the Internet in the medical field. Based on Internet technology, it deeply integrates with traditional medical services through information means, effectively connects patients, medical staff, medical equipment, and medical institutions, and realizes efficient coordination among various departments and personnel. This not only promotes the sharing and exchange of medical information and resources and improves the utilization rate of medical resources but also meets the needs of disease diagnosis and treatment of patients and improves

their medical satisfaction. At present, smart medical research mainly covers electronic medical records, medical and health big data analysis and mining, intelligent medical image analysis, intelligent assisted diagnosis, intelligent diagnosis and treatment, etc. In obstetrics, the smart medical treatment has penetrated the prediction of preeclampsia [1, 2], placenta implantation [3], type of labor [4], fetal growth [5], and other aspects. At the same time, it also made some beneficial explorations in forceps delivery [6] and treatment of obstetric complications [7, 8].

Preeclampsia is a common obstetric complication with a global prevalence of 2%–8% [9, 10]. In China, with the liberalization of fertility policy and the development of assisted reproductive technology, the number of pregnant women of advanced age, assisted reproduction, multiple pregnancies, and overweight have increased significantly, and the incidence of preeclampsia is about 2.7% [11], which is one of the major causes of maternal and perinatal death [12]. At present, the common treatment regimen for preeclampsia is symptomatic treatment with sedation, antispasmodic, antihypertensive, and correction of hypoproteinemia, and the only effective intervention plan is the termination of pregnancy. The 2020 edition of the Chinese Medical Association's guidelines for the management of hypertensive disorders in pregnancy states that, in principle, vaginal delivery should be considered for pregnant women with preeclampsia who are not severely ill and who do not have an indication for cesarean delivery. However, in actual clinical practice, many pregnant women with preeclampsia who are eligible for vaginal trial of labor directly opt for cesarean section, or because of the imperfection of medical information systems and inadequate maternal-fetal monitoring, patients with preeclampsia undergo unnecessary conversion to cesarean section during the vaginal trial of labor. This has undoubtedly increased the cesarean section rate and the incidence of surgery-related complications in our country, aggravating the economic burden on maternal families and the country.

To effectively promote vaginal delivery and reduce the rate of first cesarean section and complications of vaginal delivery, our hospital launched the construction of smart hospital in 2018 and successively introduced the integrated platform, hospital information system, laboratory information management system, hospital resource planning, remote fetal heart rate monitor, and central fetal heart rate monitor. At the same time, we established a sound medical information system and fetal monitoring system. Since 2019, relying on the above intelligent systems and equipment, our hospital started to explore the active labor process management methods under the new labor guidelines for preeclampsia pregnant women. In this study, we retrospectively analyzed the clinical data of pregnant women with preeclampsia who delivered vaginally in our hospital from January 1, 2017, to December 31, 2020, to explore the effectiveness of active labor management in improving the delivery outcomes of pregnant women with preeclampsia.

2. Materials and Methods

2.1. Materials. This is a single-center retrospective case-control study. All medical records came from Jiangxi Provincial Maternal and Child Health Hospital and were retrieved through the electronic medical record system of the hospital. The retrieval strategies were as follows: preeclampsia as the discharge diagnosis, vaginal delivery as the delivery mode, and January 1, 2017, to December 31, 2020, as the delivery time. Inclusion criteria were as follows: singleton live birth, full-term pregnancy, and spontaneous delivery. Exclusion criteria were as follows: scarred uterus, poorly controlled blood pressure, severe impairment of organ function, chronic hypertension complicated with preeclampsia, and renal hypertension complicated with antiphospholipid antibody syndrome. A total of 219 pregnant women were enrolled in this study.

2.2. Diagnostic Criteria of Preeclampsia. The diagnosis of preeclampsia refers to the Guidelines for the Diagnosis and Treatment of Hypertensive Disorders in Pregnancy (2015) formulated by the Hypertensive Diseases Group in Pregnancy of the Chinese Society of Obstetrics and Gynecology. Systolic blood pressure ≥ 140 mmHg or diastolic blood pressure ≥ 90 mmHg at or after 20 weeks of pregnancy, together with one or more of the following new-onset conditions; urinary protein ≥ 3.0 g per 24 hours, urinary protein/creatinine ratio ≥ 0.3, random urinary protein $\geq$ (+) if quantitative urine protein detection is not available; no proteinuria but with damage to vital organs such as the heart, lung, liver, kidney, or involvement of the hematological system, nervous system, digestive system, and fetus, etc.

2.3. Definition of Each Stage of Labor. (1) The first stage of labor was known as the cervical dilation period, which referred to the process from regular uterine contractions to the full dilation of the cervix; (2) the second stage of labor was also known as the fetal delivery period, referred to the process from the full dilation of the cervix to the delivery of the fetus; and (3) the third stage of labor also known as the placental delivery period referred to the process from the delivery of the fetus to the delivery of the placenta.

2.4. Key Points of Labor Management in the Control Group. The treatment of labor process was performed according to the "Expert Consensus on New Labor Process Standards and Treatment (2014)" issued by the Obstetrics and Gynecology Group of the Chinese Medical Association's Obstetrics and Gynecology Branch. The main details were as follows:

(1) Key Points of the First Stage of Labor Management. Vaginal palpation was used to assess the progress of labor, which included fetal position, fetal head position, and fetal lie. The criteria for abnormal labor were as follows: cessation of uterine orifice dilation, cessation of fetal head descent, and non-parallelism between fetal head height and uterine orifice dilation. The standard of parallelism between the height of fetal head and the dilation of the uterus is as follows,

when uterine orifice dilation size was <3 cm, 3–4 cm, 5–6 cm, 7–8 cm, and 9–10 cm, and the corresponding fetal head position was S—3~S—2, S—2~S—1, S 0~$S+1$, $S+1 \sim S+2$, and $S+2$~$S+3$. The fetal heart rate monitoring was performed every 2 hours for duration of 20–40 minutes. The observation of labor progress and fetal heart rate was carried out by the resident physician in the delivery room. If there was any abnormality during labor, the attending physician with more than 5 years of delivery room work experience must check it again. When the final judgment was abnormal labor, interventions such as sedation, manual rupture of membranes, oxytocin drip, and transfer to cesarean section were taken according to the situation.

(2) Key Points of the Second Stage of Labor Management. A bedside fetal monitor was used for continuous fetal heart rate monitoring. Doctors and midwives in the delivery room checked the fetal heart rate pattern timely to detect abnormalities as soon as possible and then actively adopt forceps or cesarean section to terminate pregnancy according to the situation.

(3) Key Points of the Third Stage of Labor Management. When the fetal shoulder was delivered, 10–20 U oxytocin was used immediately. If necessary, 100 μg carbetocin was injected intravenously after fetal delivery. If there was still active bleeding, 250 μg carboprost trometamol was added. Excluding laceration of birth canal and residual placenta, if uterine bleeding persisted after the use of a variety of uterine contraction drugs and uterine massage to promote uterine contraction, uterine tamponade hemostatic treatments would be used as soon as possible.

2.5. Key Points of Labor Management in the Study Group. During the whole labor process, all patients were provided with professional, comprehensive, and humanized Doula delivery services by systematically trained Doula personnel to give patients physical, psychological, and emotional support. The main details of labor management according to the active labor management methods under the new labor standard were as follows:

(1) Key Points of the First Stage of Labor Management. Under the premise of informed consent of patients, all pregnant women with preeclampsia were provided with intraspinal labor analgesia by an anesthesiologist in the delivery room. At the same time, life care would be strengthened, and feeding, fluid rehydration, and rest were encouraged during labor. The progress of labor was assessed by intravaginal palpation and ultrasound during labor. The model of the ultrasonic diagnostic apparatus was S8Exp, and the probe frequency was set as 2–4 mega-hertz. Transabdominal ultrasound was used to determine fetal position, and transperineal ultrasound was used to evaluate fetal head position. In transperineal ultrasound, the symphysis pubis and fetal head position of pregnant women were often used as markers. The content of the intrapartum ultrasound measurement included fetal head progression angle, midline angle, distance between fetal head and pubic symphysis, fetal head direction, and distance between fetal head and perineum. The fetal heart rate monitoring was performed by central fetal heart rate monitor, and the related parameters of fetal heart rate monitoring were the same as those of the control group. If there was any abnormality during labor, the resident physician would immediately contact the attending obstetrician who has been trained in systematic ultrasound to check again. If an abnormal labor did occur, interventions such as sedation, artificial rupture of membranes, oxytocin drip, and transfer to cesarean section were taken according to the situation.

(2) Key Points of the Second Stage of Labor Management. The central fetal heart rate monitor was used to monitor the fetal heart rate continuously throughout the whole process. Once abnormal fetal heart rate monitor pattern occurred, the central monitor would open an alarm to remind the medical staff to take active measures.

(3) Key Points of the Third Stage of Labor Management. The treatments were basically the same as the control group, except that delayed umbilical cord amputation and controlled cord pulling were encouraged when the mother and newborn were in stable condition.

2.6. Blood Pressure Management during Labor. For patients without organ dysfunction, the target blood pressure control value was 130–155/80–105 mmHg, and for those with organ dysfunction, the blood pressure control target value was 130–139/80–89 mmHg. During the first stage of labor, blood pressure was monitored every 2 hours to ensure the patient's rest and proper activities. During the second stage of labor, blood pressure was continuously monitored, and blood pressure was monitored every half hour during the third stage of labor. Of course, if blood pressure exceeded the target value during the monitoring period, the monitoring interval needed to be shortened accordingly.

2.7. Observed Indicators. Duration of 3 stages of labor, rate of forceps assisted delivery, rate of perineal laceration, rate of cervical laceration and placental abruption, postpartum hemorrhage volume within two hours, birth weight of the newborn, and Apgar score at 1 minute after birth were observed.

3. Statistical Analysis

IBM SPSS 24.0 software was used for data processing. The normally distributed data were expressed as mean ± standard deviation, the nonnormally distributed data were

expressed as median and interquartile spacing, and qualitative data were expressed as composition ratio. We used the independent-samples t test, Mann–Whitney U test, and chi-square test to analyze the statistical differences. $P < 0.05$ was considered a statistically significant difference.

4. Results

4.1. Deliveries of Pregnant Women with Preeclampsia in Our Hospital during Four Years. From 2017 to 2020, a total of 87010 deliveries were made in our hospital, including 44601 deliveries from 2017 to 2018 and 42409 deliveries from 2019 to 2020. The delivery mode and delivery volume of pregnant women with preeclampsia are shown in Table 1. Statistical analysis results showed that the number of deliveries of preeclampsia from 2019 to 2020 increased significantly compared with the previous two years, and the rate of delivery of cesarean section due to vaginal trial delivery failure decreased compared with the previous two years.

4.2. Clinical Characteristics of Mothers and Newborns. The age distribution of patients with preeclampsia was 17~45 years, the gestational week of delivery was 259~291 days, and the number of pregnancies ranged from 1 to 6, and these in the control group were 18~43 years, 259~ 293 days, and 1~8 times. In addition, there were 50 (39.37%) women in the study group and 32 (34.78%) women in the control group who had histories of childbirth. There were no differences in the demographic parameters of mothers and newborns between the two groups, as shown in Table 2 for details.

4.3. Vaginal Delivery Outcomes of the Two Groups. The vaginal delivery outcomes are compared between the two groups of preeclampsia women. We found that the duration of the first and second stages of labor was slightly shorter in the study group than in the control group, and the rates of episiotomy and postpartum hemorrhage volume within two hours were lower than those in the control group, but the duration of the third stage of labor, rate of forceps delivery, rate of cervical laceration, and placental abruption were not significantly different from those in the control group. See Table 3.

4.4. Comparison of the Causes of Postpartum Hemorrhage between the two Groups. A total of 39 patients in the two groups suffered from postpartum hemorrhage. The main causes of postpartum hemorrhage were uterine weakness, followed by placental factors and birth canal lacerations. No postpartum hemorrhage caused by coagulation dysfunction occurred. Statistical analysis showed that there was no difference in the causes of postpartum hemorrhage between the two groups. See Table 4.

4.5. Comparison of Hemostatic Methods between the Two Groups. The hemostatic measures of the two groups were mainly the combination of two pro-uterine contraction drugs, followed by a single pro-uterine contraction drug, and in a few cases, three contraction drugs or even uterine tamponade were used. Statistical analysis showed no difference in the methods of hemostasis between the two groups. The detailed results are shown in Table 5.

5. Discussion

In the context of rising cesarean section rate globally, according to the statistics of National Maternal and Child Health Care, the cesarean section rate in China increased from 28.80% to 36.70% from 2008 to 2018 [13], which seriously exceeded the alarm level of the cesarean delivery rate set by the World Health Organization [14]. At the same time, a series of serious complications caused by cesarean section, such as placenta previa, uterine incision pregnancy, placenta implantation, and uterine rupture, have seriously affected maternal safety and also impose heavy economic burden to the society. In recent years, vaginal delivery has been advocated worldwide to reduce the primary cesarean section rate [15], and in 2021, China's Health and Welfare Commission listed "reducing the incidence of vaginal delivery complications" as an improvement goal of medical quality and safety. As a common obstetric complication, preeclampsia affects a large number of women in China. It was not difficult to find that the incidence of simple preeclampsia in our hospital increased significantly in the past two years, and the number of pregnant women who directly chose cesarean section for various reasons also increased under the background of the continuous improvement of medical technology, which was contrary to the goal of improving medical quality and safety in China. Therefore, based on the safety of mother and baby, it is one of the key works of obstetricians to promote vaginal delivery of pregnant women with preeclampsia.

The introduction of the smart medical information system is one of the major initiatives to promote vaginal delivery in our hospital. In 2019, the hospital completed the interconnection of the outpatient electronic case system, remote fetal heart rate monitoring system, inpatient electronic case system, nursing system, testing system, and imaging system so that obstetricians could get a comprehensive and detailed understanding of patients' blood pressure fluctuations during pregnancy when receiving pregnant women with preeclampsia. Changes in blood test indicators, drug use, and fetal growth indicators are conducive to the development of detailed prenatal examination plans and reasonable delivery plans for patients and reduce misdiagnosis and missed diagnosis caused by incomplete and accurate patient history review. When the patients were sent to the delivery room to wait for delivery, the delivery room medical staff could also achieve good communication with outpatient doctors, ward doctors, and nurses through the medical information system, better observe the progress of labor, and give timely and accurate intervention measures according to the timely feedback of the central fetal heart rate monitor.

Intrapartum ultrasound is a technique of labor monitoring that has become popular in clinic in recent years. The monitoring includes fetal position, fetal head position, and

Table 1: Preeclampsia deliveries in our hospital during four years.

	01/01/2017–31/12/2018	01/01/2019–31/12/2020	P value	Method
Total number of deliveries (cases)	44601	42409		
Preeclampsia deliveries (cases/rate)	261 (0.59%)	440 (1.04%)	< 0.001	Pearson's chi-square
Vaginal deliveries (cases/rate)	92 (35.25%)	127 (28.86%)	0.078	Pearson's chi-square
Cesarean delivery (cases/rate)	169 (64.75%)	313 (71.14%)	0.078	Pearson's chi-square
Births converted to cesarean section (cases/rate)	29 (17.16%)	33 (10.86%)	0.038	Pearson's chi-square

Table 2: Clinical characteristics of mothers and newborns of study population.

	Study group ($n = 127$)	Control group ($n = 92$)	P value	Method
Age (years)	28.65 ± 5.32	29.36 ± 5.62	0.340	Independent-samples t
Gestational age (days)	273.85 ± 7.68	275.47 ± 8.26	0.138	Independent-samples t
Gravidity (times)	2 (1, 3)	2 (1, 3)	0.598	Mann–Whitney
Parity (times)	1 (1, 2)	1 (1, 2)	0.520	Mann–Whitney
Neonatal weight (kg)	3.29 ± 0.42	3.29 ± 0.39	0.941	Independent-samples t
Apgar score at 1 minute after birth	10 (10, 10)	10 (10, 10)	0.091	Mann–Whitney

Data are presented as median (interquartile spacing) or mean ± standard deviation.

Table 3: Vaginal delivery outcomes of these study groups.

	Study group ($n = 127$)	Control group ($n = 92$)	P value	Method
Duration of the first stage of labor (minutes)	440 (245, 600)	480 (280, 750)	0.496	Mann–Whitney
Duration of the second stage of labor (minutes)	29 (16, 48)	31 (15, 60)	0.787	Mann–Whitney
Duration of the third stage of labor (minutes)	8 (5, 10)	8 (5, 10)	0.759	Mann–Whitney
Episiotomy (cases/rate)	39 (30.71%)	41 (44.57%)	0.042	Pearson's chi-square
Forceps delivery (cases/rate)	4 (3.15%)	4 (4.35%)	0.919	Continuity correction
Perineal laceration (cases/rate)	32 (25.20%)	24 (26.09%)	0.882	Pearson's chi-square
Cervical laceration (cases/rate)	7 (5.51%)	1 (1.09%)	0.174	Continuity correction
Postpartum hemorrhage volume within two hours (mL)	280 (200, 355)	325 (240, 433.75)	0.037	Mann–Whitney

Data are presented as median (interquartile spacing) or mean ± standard deviation.

Table 4: Comparison of causes of postpartum hemorrhage between the two study groups.

	Weak contractions (cases/rate)	Placental factors (cases/rate)	Birth canal laceration (cases/rate)
Study group ($n = 15$)	10 (66.67%)	3 (20.00%)	2 (13.33%)
Control group ($n = 24$)	18 (75.00%)	4 (16.67%)	2 (8.33%)
P value		1.000	
Method		Monte Carlo	

Table 5: Comparison of hemostatic methods between the two groups.

	One uterine contraction drug (cases/rate)	Two uterine contraction drugs (cases/rate)	Three uterine contraction drugs (cases/rate)	Uterine tamponade (cases/rate)
Study group	59 (46.46%)	62 (48.82%)	5 (3.94%)	1 (0.79%)
Control group	39 (42.39%)	48 (52.17%)	4 (4.35%)	1 (1.09%)
P value		0.932		
Method		Monte Carlo		

fetal lie. Commonly used parameters include fetal head progression angle, midline angle, distance between fetal head and pubic symphysis, fetal head direction, and distance between fetal head and perineum. With the guidance of ultrasound imaging, we can quantify each parameter index to evaluate the progress of labor more accurately and to seize the right time to intervene in labor to avoid unnecessary vaginal-assisted labor and intermediate cesarean delivery. In the past two years, we have adopted the method of vaginal palpation combined with intrapartum ultrasound during labor to comprehensively evaluate the labor process. On the one hand, it avoided the misjudgment of labor process caused by inadequate clinical experience in obstetrics and inaccurate palpation of vagina, and on the other hand, it avoided the measurement error of parameters caused by incomplete ultrasound knowledge. Compared with the outcomes of the two groups, the duration of the first and second stages of labor in the study group was slightly shorter than that in the control group, and the rates of forceps delivery and perineal laceration were also lower than in the

control group, while the rate of cesarean section and perineal lateral resection was significantly lower than that in the control group. The results of this study gave positive implications for the observation of labor progress of pregnant women with preeclampsia by vaginal palpation combined with intrapartum ultrasound.

While efforts are being made to promote vaginal delivery in preeclampsia, we need to be more proactive in preventing and reducing the delivery complications of preeclampsia. Pregnant women with preeclampsia are often accompanied by decreased proteinuria and plasma albumin levels and are prone to edema of the limbs and skin, as well as edema of the myometrium, which can lead to weak uterine contractions. Combined with the use of sedatives and antispasmodics during labor, the probability of postpartum hemorrhage due to weak uterine contractions will be higher. In this study, the main cause of bleeding in both groups was uterine contraction weakness. The median postpartum hemorrhage volume within two hours in the study group was 280 mL and the maximum bleeding volume was 1650 mL, while in the control group, they were respective 325 mL and 1500 mL. On the basis of the same uterine pathology in preeclamptic women, the age, gestational age, number of gravidities, parities and fetuses, causes of bleeding, and hemostatic measures in both two groups were basically the same, and the postpartum hemorrhage volume within two hours was less in the study group than in the control group. The possible reasons were considered as follows: (1) the extensive development of Doula delivery and labor analgesia services could relieve the tension of pregnant women with preeclampsia, avoid the drastic fluctuation of blood pressure during labor, and promote the changes in norepinephrine and other endocrine hormones to strengthen the contractile force of the uterus. (2) Warm life care, appropriate food, and fluid supplementation ensured adequate energy supply during labor, which not only reduced the failure of vaginal delivery due to insufficient maternal blood volume but also enhanced uterine contraction.

6. Summary

Under the environment of actively promoting vaginal delivery, reducing the rate of first cesarean section and the complication rate of vaginal delivery, actively promoting the construction of smart hospital, improving medical information system, ensuring convenient and adequate maternal and fetal monitoring equipment, and improving the quality of delivery services can not only effectively reduce the probability of failed vaginal trial of labor and perineal scoliosis in preeclamptic women but also effectively reduce the postpartum hemorrhage volume within two hours without increasing the use of hemostatic drugs. As obstetricians, we should strengthen our theoretical knowledge and skills in obstetrics, as well as our knowledge of psychology and ultrasound imaging, to deal with various emergencies that may occur during labor.

Data Availability

The labeled dataset during this study is available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declared that no potential conflicts of interest exist.

Authors' Contributions

Siming Xin and Xianxian Liu contributed equally to this work.

Acknowledgments

This work was supported by the Science and Technology Project of Jiangxi Province (grant nos. 20192BBGL70003, 20203BBGL73130, and 20212BAB216065).

References

[1] F. Xie, T. Im, and D. Getahun, "A computerized algorithm to capture patient's past preeclampsia and eclampsia history from prenatal clinical notes," *Health Informatics Journal*, vol. 25, no. 4, pp. 1299–1313, 2019.

[2] S. Luo, N. Cao, Y. Tang, and W. Gu, "Identification of key microRNAs and genes in preeclampsia by bioinformatics analysis," *PLoS One*, vol. 12, no. 6, Article ID e0178549, 2017.

[3] H. Sun, H. Qu, L. Chen et al., "Identification of suspicious invasive placentation based on clinical MRI data using textural features and automated machine learning," *European Radiology*, vol. 29, no. 11, pp. 6152–6162, 2019.

[4] J. Alberola-Rubio, J. Garcia-Casado, G. Prats-Boluda et al., "Prediction of labor onset type: spontaneous vs induced; role of electrohysterography?" *Computer Methods and Programs in Biomedicine*, vol. 144, pp. 127–133, 2017.

[5] A. I. Naimi, R. W. Platt, and J. C. Larkin, "Machine learning for fetal growth prediction," *Epidemiology*, vol. 29, no. 2, pp. 290–298, 2018.

[6] S. Xin, Z. Wang, H. Lai et al., "Clinical effects of form-based management of forceps delivery under intelligent medical model," *Journal of Healthcare Engineering*, vol. 2021, Article ID 9947255, 2021.

[7] A. Seitinger, A. Rappelsberger, H. Leitich, M. Binder, and K.-P. Adlassnig, "Executable medical guidelines with arden syntax-applications in dermatology and obstetrics," *Artificial Intelligence in Medicine*, vol. 92, pp. 71–81, 2018.

[8] S. Barbounaki, A. Sarantaki, and K. Gourounti, "Fuzzy logic intelligent systems and methods in midwifery and obstetrics," *Acta Informatica Medica*, vol. 29, no. 3, pp. 210–215, 2021.

[9] E. Pierik, J. R. Prins, H. van Goor et al., "Dysregulation of complement activation and placental dysfunction: a potential target to treat preeclampsia?" *Frontiers in Immunology*, vol. 10, p. 3098, 2019.

[10] S. Rana, E. Lemoine, J. P. Granger, and S. A. Karumanchi, "Preeclampsia," *Circulation Research*, vol. 124, no. 7, pp. 1094–1112, 2019.

[11] Y.-C. Zhu, H.-X. Yang, Y.-M. Wei et al., "Analysis of correlation factors and pregnancy outcomes of hypertensive disorders of pregnancy - a secondary analysis of a random sampling in beijing, China," *Journal of Maternal-Fetal and Neonatal Medicine*, vol. 30, no. 6, pp. 751–754, 2017.

[12] V. L. Bilano, E. Ota, T. Ganchimeg, R. Mori, and J. P. Souza, "Risk factors of pre-eclampsia/eclampsia and its adverse outcomes in low- and middle-income countries: a WHO

secondary analysis," *PLoS One*, vol. 9, no. 3, Article ID e91198, 2014.

[13] H.-t. Li, S. Hellerstein, Y.-b. Zhou, J.-m. Liu, and J. Blustein, "Trends in cesarean delivery rates in China, 2008-2018," *JAMA*, vol. 323, no. 1, pp. 89–91, 2020.

[14] B. Chalmers, "WHO appropriate technology for birth revisited," *BJOG: An International Journal of Obstetrics and Gynaecology*, vol. 99, no. 9, pp. 709-710, 1992.

[15] A. D. Bell, S. Joy, S. Gullo, R. Higgins, and E. Stevenson, "Implementing a systematic approach to reduce cesarean birth rates in nulliparous women," *Obstetrics & Gynecology*, vol. 130, no. 5, pp. 1082–1089, 2017.

Hindawi
Journal of Healthcare Engineering
Volume 2022, Article ID 5707930, 11 pages
https://doi.org/10.1155/2022/5707930

Research Article

Automated Facial Expression Recognition Framework Using Deep Learning

Saad Saeed,[1] Asghar Ali Shah,[1] Muhammad Khurram Ehsan,[1] Muhammad Rizwan Amirzada,[2] Asad Mahmood,[3] and Teweldebrhan Mezgebo[4]

[1]*Faculty of Engineering, Bahria University, Lahore Campus, Lahore, Pakistan*
[2]*Faculty of Engineering and Computer Science, National University of Modern Languages, Islamabad, Pakistan*
[3]*Department of Electrical and Computer Engineering, Comsats University, Islamabad, Wah Campus, Pakistan*
[4]*Ethio Telecom, Addis Ababa, Ethiopia*

Correspondence should be addressed to Teweldebrhan Mezgebo; tewelde2000@gmail.com

Received 6 February 2022; Accepted 15 March 2022; Published 31 March 2022

Academic Editor: Basem M. Elhalawany

Facial expression is one of the most significant elements which can tell us about the mental state of any person. A human can convey approximately 55% of information nonverbally and the remaining almost 45% through verbal communication. Automatic facial expression recognition is presently one of the most difficult tasks in the computer science field. Applications of facial expression recognition (FER) are not just limited to understanding human behavior and monitoring person's mood and the mental state of humans. It is also penetrating into other fields such as criminology, holographic, smart healthcare systems, security systems, education, robotics, entertainment, and stress detection. Currently, facial expressions are playing an important role in medical sciences, particularly helping the patients with bipolar disease, whose mood changes very frequently. In this study, an algorithm, automated framework for facial detection using a convolutional neural network (FD-CNN) is proposed with four convolution layers and two hidden layers to improve accuracy. An extended Cohn-Kanade (CK+) dataset is used that includes facial images of different males and females with expressions such as anger, fear, disgust, contempt, neutral, happy, sad, and surprise. In this study, FD-CNN is performed in three major steps that include preprocessing, feature extraction, and classification. By using this proposed method, an accuracy of 94% is obtained in FER. In order to validate the proposed algorithm, K-fold cross-validation is performed. After validation, sensitivity and specificity are calculated which are 94.02% and 99.14%, respectively. Furthermore, the f1 score, recall, and precision are calculated to validate the quality of the model which is 84.07%, 78.22%, and 94.09%, respectively.

1. Introduction

A facial expression is to be accounted to know about the emotional state, psychopathology, cognitive activity, and intention of an individual. In interpersonal relations, the facial expressions play an expressive and communicative role. The importance of facial expression recognition (FER) can be gauged from the fact that it can describe any person's mental state or mood. Its applications are not just limited to understand human behaviour, viewing a person's mood, or judging the mental state of humans. It is also penetrating into other fields such as criminology, holographic, smart healthcare systems, security systems, education, robotics, entertainment, multimedia communication, and stress detection [1–7]. The integration of facial expressions in these fields shows that facial expressions have an essential part in human life. Automatic FER is presently one of the most challenging tasks in the field of computer sciences. Expressions can be conveyed through gestures and communications. It does not just depend on the human face. Mehrabian et al. [8] say that a person can transfer only 7% of the information context orally, while 38% can be transferred through voice tone, rhythm, and how speedily or slowly a person speaks. On the other hand, the information which is

transferred through facial expressions is 55%. Through facial expression, one can understand the mental state of the other person. Ekman et al. [9, 10] introduced the facial emotions into seven basic categories which include happy, fear, sad, surprise, anger, contempt, and disgust.

Applications of facial expressions cover a huge area of our society and are not just limited to some specific fields. In medical sciences, FER is beneficial for bipolar patients. As described in [11], doctors are trying to detect and monitor patients' behavior, such as how a bipolar patient feels and how they behave during their disease. In [12], an intelligent FER system is devised such that facial images are given as input and the system can detect the expressions of a human face. There are in total 8 expressions [13, 14] that a human can express which include fear, happy, surprise, neutral, angry, contempt, sad, and disgust state.

In this study, the eight expressions are used, and an algorithm named FD-CNN is proposed to improve the accuracy of FER. Three major building blocks: preprocessing, feature extraction, and classification are further categorized to obtain the underlying objective. In this process of FER, the extended Cohn-Kanade (CK+) dataset is chosen. It has images extracted from 123 different people, including males and females. All images are captured from the front angle and divided into eight different categories. In the first phase of preprocessing, all the images are reshaped to 150×150 pixels so that all images have an equal size. Further, in preprocessing, images are randomly rotated between 0 and 180 degrees and zoomed. Images are also flipped horizontally and vertically. In the next phase, images are further processed to extract features. In feature extraction, a filter or kernel is applied to the image, and this kernel can be of different sizes according to the required features. In this study, 3×3 kernel is used to address the tradeoff between the feature extraction requirements and computation cost. After applying the said sized kernel, the output comes out in the form of face curves and edges. Once the features are extracted, then there is a need to keep only useful features, and max pooling helps out in this regard. Classification is the third phase of the proposed methodology, and it is responsible for detecting the correct labels. In classification, fully connected layers are used, and these fully connected layers use two hidden layers further. Each hidden layer has multiple nodes that have some weights. There is a process called forward propagation in which the weight to these nodes are multiplied with bias values and then summed up all. In backpropagation, the weight of the hidden layer's nodes is adjusted, and the model can find the true label of an input image.

The remaining part of this paper is organized as follows: in Section 2, the state-of-the-art-work of FER is described. Section 3 discusses the methodology of the proposed algorithm in detail. Simulations and results are described in Section 4. In the end, the presented work is concluded in Section 5.

2. Literature Review

Facial expressions play a crucial role in our everyday life while communicating with others. In 1978, the first ever automatic FER system was introduced by Suwa [15], in which he explained how he identified 20 spots in an image sequence to find facial expressions.

Pantic and Rothkrantz [16] proposed a system for automatic recognition of facial gestures from a static coloured face image. They employed the multidetector technique for facial feature localization to combine facial regions' contours and profile contours like eyes and mouth. Using rule-based reasoning, the author extracted 10 fiducial points of the profile contour and 19 fiducial points of the facial component contours, while 32 facial muscle action units [17] were recognized with 86% accuracy.

Chen et al. [18] proposed a system for facial expression using a hybrid method based on geometric and appearance features that found the difference of points from a neutral facial image and an emotional facial image. This hybrid feature contains the local texture difference and the facial feature point displacement. The accuracy of the proposed method is 95% on the CK + database by using the support vector machine (SVM) classifier.

Uçar et al. [19] proposed a methodology for FER by using the curvelet transform (CT) and the online sequential extreme learning machine (OSELM) with a radial basis function. In the first step, the face is divided into some small regions which are called local regions, and then the CT is smeared to that local region of the face. The main purpose of doing this is to reduce the curvelet coefficients so that they can be easily classified within a minimum time. Furthermore, they calculated the standard deviation, entropy, and mean of curvelet coefficient to generate the set of features for each region. To increase the classification and reduce the time required in choosing the hidden node number, they used the spherical clustering method on the feature set. The learning is then called OSELM-SC and it consists of two parts. The first part checks the performance of OSELM-SC on different datasets, and the second part tests the accuracy of the FER algorithm on the Japanese Female Facial Expressions (JAFFE) and Cohn Kanade (CK) database.

One of the challenging tasks for accurate FER is the extraction of emotional features correctly from the input images. In [20], salient distance features are used to solve the problem of extracting facial features accurately, while these salient features are obtained by extracting patch-based 3D Gabor features. The results show that the good CRR, i.e., the correct recognition rate and the system performance, is improved by considering the facial and muscle movements. They get a good accuracy rate by using the JAFFE and CK databases.

In [21], a kernel-based manifold learning method is proposed, which is called KDIsomap, i.e., kernel discriminant isometric mapping. KDIsomap extracts the discriminant information nonlinearly. It is also used as a dimensionality reduction on extracted facial features. They use the closest neighbor classifier to the Euclidean metric to recognize facial expressions. This classifier is applied to two publicly available databases, which are JAFFE and CK. Using KDIsomap, the obtained accuracy is 94.88% on the CK database and 81.59% accuracy is obtained on the JAFFE database.

In human-computer interaction, FER plays a crucial role. Jabid et al. [22] presented the appearance-based feature descriptor, which is, i.e., local directional pattern (LDP). This LDP represents the facial geometry and evaluates FER performance. In each pixel, by computing the edge response values, LDP features are obtained. These features are then encoded into an 8-bit binary number. Principal component analysis (PCA) and AdaBoost are then used for dimensionality reduction, which in turn benefits us with less computation and better classification accuracy. Template matching and SVM are two machine learning methods that are applied to CK and JAFFE databases to recognize facial expressions.

Negative emotional states can cause a mental health problem. So, to improve people's health, especially for the elderly, Uddin et al. [23] proposed an efficient method for the FER which is helpful in emotional healthcare systems. In the proposed FER, a video data is used to detect the facial states of human, also the feature extraction is employed to extract the predominant features from the given dataset. For robust feature extraction, the local directional position pattern (LDPP) method is used. Furthermore, they used PCA and generalized discriminant analysis (GDA), respectively, to reduce the selected features, and these features were then used in the deep belief network (DBN) for the FER.

In [24], they used an extensive range of head postures for FER. Their methodology, mapped the nonfrontal view features with the corresponding features of the frontal view within the same facial expression. For FER, they first estimated the posture of the head in the given input image and then applied mapping. The extracted features after mapping were very trustworthy for recognition purposes. The authors evaluated their results by applying the methods on multi-PIE and BU3DFE datasets, and good results were achieved.

In [25], the authors tried to do local learning for FER in which they used a convolutional neural network (CNN) for automatic feature learning and a bag of visual words (BOVW) model is used for handcrafted feature learning. These two types of features are combined to find out the best result of FER. There are three steps for the local learning framework which are used in the paper. First, K nearest neighbour (KNN) is used to select only those images from the training data that are closest to the input test image. Secondly, SVM is trained using selected images and then, based on that training, SVM classifies the input test images efficiently. In this research, the authors used 4 different datasets which are FER 2013, FER+, AffectNet with eight expressions, and AffectNet with seven expressions. The accuracies of the proposed method with said datasets are 75.42%, 87.76%, 59.58%, and 63.31%, respectively.

In [26], automatic FER is performed by using occlusion and pose variation. The authors used the real-world FER dataset with different poses and occlusion attributes, and the method that was used is region attention network (RAN). They validated RAN using four different datasets named as SFEW, RAF-DB, AffectNet, and FER Plus. The obtained accuracies are 56.4%, 86.9%, 59.5%, and 89.16%, respectively.

3. Dataset

In this study, the dataset used is CK + as detailed in Table 1. This dataset is obtained from the Kaggle dataset website. The images consist of 640×490 pixels from 123 different human faces including males and females.

There are 920 images in total which are used in this research for training and testing of the proposed FD-CNN. These images are labeled with eight emotions including fear, happy, sad, contempt, disgust, neutral, surprise, and anger, as shown in Figure 1. All these images are captured from the front angle.

4. Methodology

The methodological framework is pictorially presented in Figure 2. The methodology consists of preprocessing, feature extraction, and classification that are discussed as following.

Figure 2 shows that the acquired dataset is preprocessed at the first phase of the proposed framework FD-CNN. This FD-CNN uses CNN with four convolutional layers and two hidden layers. Once preprocessing is completed, information is passed to the next phase called feature extraction. In the feature extraction phases, many substages help to extract features, and only valuable features are kept for further processing. In the end, classification is performed to recognize label of the input image correctly. To find the correct label of the input image, forward and backward propagation are exploited for error estimation and minimization.

4.1. Preprocessing. Preprocessing is used to improve the system performance by transforming the inconsistent dataset into a consistent dataset. Consistency is obtained by reshaping all the images into 150×150 pixels. Further, in this step, all images are randomly rotated between 0 and 180 degrees, and they are also zoomed randomly. Images are also flipped horizontally and vertically too.

4.2. Architecture of CNN. CNN is one of the conventional forms of deep neural networks. It takes images as input, and with the help of learnable weights and biases, it can differentiate in different images. These weights and biases are used in hidden layers. CNN is mostly used for analyzing visual imagery [27, 28].

Main building blocks of CNN include feature extraction and other one is classification. Feature extraction is further subdivided into convolution, padding, nonlinearity, and pooling. In CNN, classification contains a fully connected layer [29], as shown in Figure 3. This fully connected layer is made up of hidden layers, and one hidden layer contains many neurons. Each neuron in a layer is linked to all the neurons in the next hidden layer [30].

4.3. Feature Extraction. Feature extraction is performed to find out the useful information which is exploited for the classification of images. For example, if we have an image of a human face, then through feature extraction, we can find out the eyes, nose, and lips of the input human face [31].

Table 1: Metadata for dataset.

Sr. No	Emotions	Number of emotions
1	Anger (An)	45
2	Contempt (Co)	18
3	Disgust (Di)	59
4	Fear (Fe)	25
5	Happy (Ha)	69
6	Neutral (Nu)	593
7	Sad (Sa)	28
8	Surprise (Su)	83

Figure 1: Eight emotions of CK + dataset.

Feature extraction has further substages which are explained as following.

4.3.1. Convolution. Convolution is performed on an input image by using a filter or a kernel in CNN. The kernel used in this paper is 5×5 in size. In order to perform this filtering and convolution, we do a scan of the whole image [32]. We start scanning from the top left side of the image and go to the right side and then move one bit down and repeat the same procedure from left to right until the whole image is scanned.

4.3.2. Padding. Padding can be categorized as valid padding or same padding [33]. In valid padding, the dimensionality is to be reduced from the original one, while in same padding, the dimensionality is either to be increased or remain the same as compared to the original one, depending on the case or situation [34]. In the proposed FD-CNN, the same padding is used, in which the dimensionality is to remain same.

4.3.3. Nonlinearity. After moving our filter from top left to right on our original input image, the resultant output will pass through one more step called a nonlinearity function or activation function. This activation function is called ReLu the Rectified Linear Unit and it is mainly used in CNN for feature extraction [35]. The working of this ReLu function is so simple. It just converts all the negative values of the matrix into 0 and does not change the positive values.

4.3.4. Pooling. Once you get the feature map from the convolutional layer, it is time to add a subsampling layer or pooling in CNN layers. Just like the convolutional layer, the pooling layer also reduces the size of the convolved features [36]. After the dimensionality reduction, there is less computational power required to process the data. Pooling is very helpful for extracting features from those images which are rotational and positional invariant. Pooling is also helpful in shortening the training time and controlling overfitting.

There are two types of pooling which can be applied [37]. The first one is max pooling, and the other one is average pooling. In max pooling, it returns the maximum value from a region where the kernel or filter overlaps with the input image. In average pooling, the average of the values from the region where kernel overlaps with the input image is returned. In this FD-CNN, we used max pooling. The reason behind choosing it is that max pooling avoids noise from the image in the calculation because it takes the maximum value from the selected region, while in average pooling noise can be added because it takes the average of the whole region.

4.3.5. Fully Connected Layer. In the CNN, there are two hidden layers, and each hidden layer contains different sets of neurons. Each neuron is fully connected with the neurons of the next layer [38]. In this network, the information comes from the feature extraction and is used in the hidden layer where the output values come from the multiplication of each neuron weight with bias value [39], then the summation of these obtained values is compared with the actual output value, and this process is called forward propagation [40]. If our calculated value is not nearest to the actual value, then the weights of neurons are adjusted, and this process is called backpropagation [41]. In the end, when the required output value is gained by using backpropagation, the model then classifies the input images and labels them according to the obtained values.

5. Simulation and Results

The proposed model implementation is also validated by using different tests. These tests include independent tests, K-fold cross-validation, the confusion matrix, recall, precision, F1 measure, and the ROC curve. This study calculates the system's sensitivity and specificity, and the results are to be discussed.

5.1. Evaluation. In order to validate the performance of FD-CNN, 10-fold cross validation test is performed. The confusion matrix, sensitivity, and specificity of the FD-CNN are then evaluated.

5.2. Independent Test Set. In the independent test, the dataset is segregated into training and testing by an 8 : 2 split ratio, but this split ratio can be changed according to the needs. After that, the proposed model is trained using a training dataset and the model is validated using a test dataset. The CNN classifier is used for the classification of images into eight categories, and the accuracy of FD-CCN is evaluated.

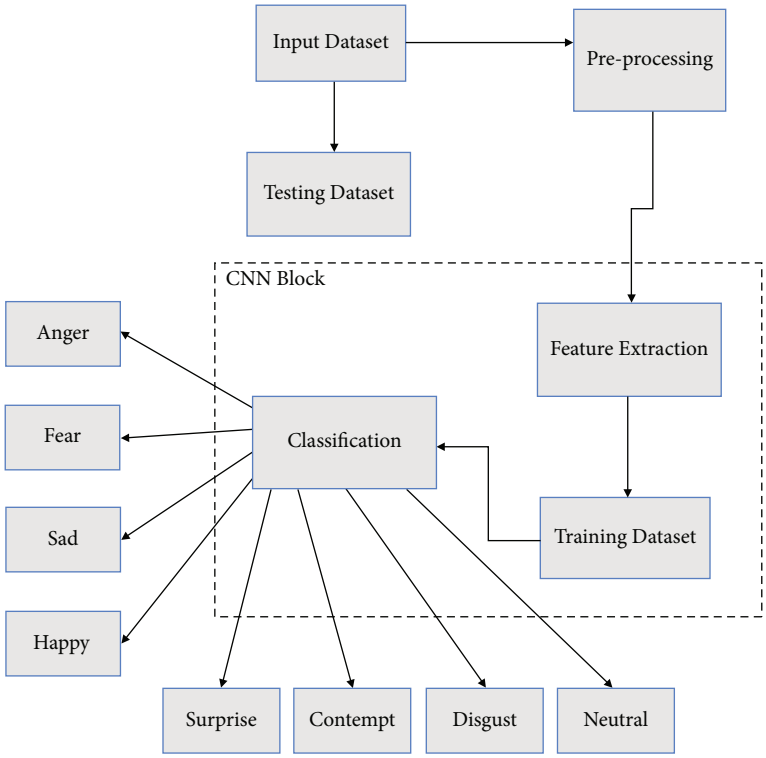


Figure 2: Methodology.

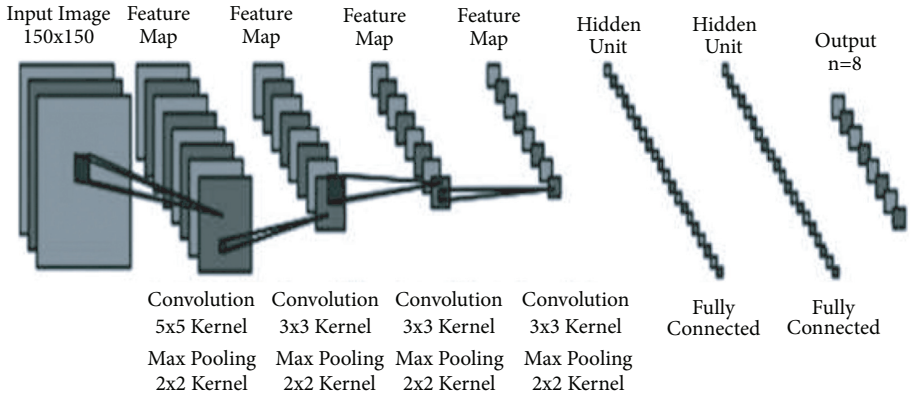


Figure 3: Architecture of the convolutional neural network.

Besides the precision, the recall f1 score, sensitivity, and specificity are calculated.

5.2.1. Confusion Matrix. The CK + dataset is used to train and validate the FD-CNN. A confusion matrix of the FD-CNN is obtained, as shown in Table 2. It contains the eight classes of facial emotions that include surprise, sad, neutral, happy, fear, disgust, contempt, and anger.

In Table 2, the rows show the actual classes, and the columns show the predicted classes. According to the confusion matrix, the results describe that the model efficiently predicts the classes of disgust, happy, neutral, and surprise with high accuracy of 82%, 100%, 100%, and 89%, respectively. Anger, contempt, and fear classes are predicted with 29%, 33%, and 25% accuracy, respectively. In contrast, a sad class is predicted with satisfactory results. The confusion matrix also shows that the anger class is misclassified as 43% disgust and 29% neutral. The contempt class is 67% misclassified as neutral. Disgust and fear are also misclassified. In the case of disgust, it is misclassified as 12% anger and 6% neutral, whereas fear is misclassified as 25% disgust and 50% neutral. Similarly, sad and surprise are misclassified. In the case of sad it is misclassified as 17% anger, 17% contempt, 33% neutral, and 17% surprise whereas surprise is misclassified as 11% neutral only.

In Table 3, all the true and predicted values of confusion matrix for 100 epochs with data duplication are shown.

Table 2: Confusion matrix for 100 epochs.

	An	Co	Di	Fe	Ha	Ne	Sa	Su
An	29%	0	43%	0	0	29%	0	0
Co	0	33%	0	0	0	67%	0	0
Di	12%	0	82%	0	0	6%	0	0
Fe	0	0	25%	25%	0	50%	0	0
Ha	0	0	0	0	100%	0	0	0
Ne	0	0	0	0	0	100%	0	0
Sa	17%	17%	0	0	0	33%	17%	17%
Su	0	0	0	0	0	11%	0	89%

Table 3: Confusion matrix for 100 epochs with data duplication.

	An	Co	Di	Fe	Ha	Ne	Sa	Su
An	94%	0	6%	0	0	0	0	0
Co	0	94%	0	0	0	6%	0	0
Di	0	0	100%	0	0	0	0	0
Fe	0	0	0	90%	0	10%	0	0
Ha	0	0	0	0	100%	0	0	0
Ne	0	0	0	0	0	100%	0	0
Sa	0	0	0	0	0	14%	86%	0
Su	0	0	0	0	0	0	0	100%

Table 4: Confusion matrix for 1000 epochs.

	An	Co	Di	Fe	Ha	Ne	Sa	Su
An	71%	0	14%	0	0	14%	0	0
Co	0	67%	0	0	0	33%	0	0
Di	0	0	94%	0	0	6%	0	0
Fe	0	0	25%	50%	0	25%	0	0
Ha	0	0	0	0	100%	0	0	0
Ne	0	0	0	0	0	99%	1%	0
Sa	0	0	0	0	0	33%	50%	17%
Su	0	0	0	0	0	6%	0	94%

Data duplication increases the size of the dataset, and with the increase of the dataset, the model can train itself better. Table 3 shows eight rows and eight columns of emotion, while the rows show the actual classes and the columns show the predicted classes. According to the confusion matrix, the results validate that the model efficiently predicts the classes of anger, contempt, disgust, fear, happiness, neutral, and surprise with high accuracy of 94%, 94%, 100%, 90%, 100%, 100%, and 100%, respectively. Anger and contempt classes are predicted with 94% and 94% accuracy, but misclassified with 6% disgust in the case of anger and 6% neutral in the case of contempt. In comparison, fear classes are predicted at 90% but misclassified with 10% neutral class. The sad class is only predicted 86% correct and easily misclassified with 14% neutral class.

Table 4 shows all the true and predicted values of confusion matrix for 1000 epochs.

According to the confusion matrix in Table 4, the results describe that the model efficiently predicts the classes of disgust, happiness, neutral, and surprise with high accuracy of 94%, 100%, 99%, and 94%, respectively. Anger and contempt classes are predicted with an accuracy of 71% and 67%, respectively. At the same time, fear and sad classes are predicted with satisfactory results. The confusion matrix also shows that the anger class is misclassified as 14% disgust and 14% neutral. Contempt is misclassified as 33% neutral. Disgust and fear are also misclassified; in the case of disgust, it is misclassified as 6% neutral, whereas fear is misclassified as 25% disgust and 25% neutral. Similarly, sad and surprise are misclassified; in the case of sad, it is misclassified as 33% neutral and 17% surprise, whereas it is misclassified as 6% neutral only in the case of a surprise. The neutral class is just 1% misclassified as sad. The low accuracy of fear and sadness compared to others is due to the smaller number of images of these classes. As the information is not well sufficient to train the model, accuracy will also be compromised.

5.2.2. Model Loss. A model loss is a metric to penalize the model if it does not correctly predict the input data. The loss of the proposed FD-CNN is shown for 100 epochs without data duplication in Figure 4. While in Figure 5, loss of the FD-CNN is shown for 100 epochs with data duplication.

Figure 6 shows the FD-CNN loss for 1000 epochs without data duplication.

5.2.3. Model Accuracy. The training and testing accuracy of FD-CNN for 100 epochs is shown in Figure 7 without data duplication. While the accuracy of FD-CNN for 100 epochs with data duplication is shown in Figure 8.

Figure 9 shows the FD-CNN accuracy for 1000 epochs without data duplication.

In order to validate the performance of the proposed methodology, a 10-fold cross validation is also performed. In 10-fold cross validation, the dataset is shuffled randomly and then divided into ten equal folds. In every iteration, a unique fold is taken as test data and the remaining folds as training data. After training, FD-CNN is validated using a test fold with a certain accuracy. When all 10 folds are performed, the obtained accuracies are summed up and their average is taken to calculate the final score of the accuracy, as shown in Table 5.

5.2.4. Evaluation through Precision, Recall, and F1 Score. The quality of a proposed model FD-CNN is to be evaluated with measured recall, precision, and f1 score. These metrics

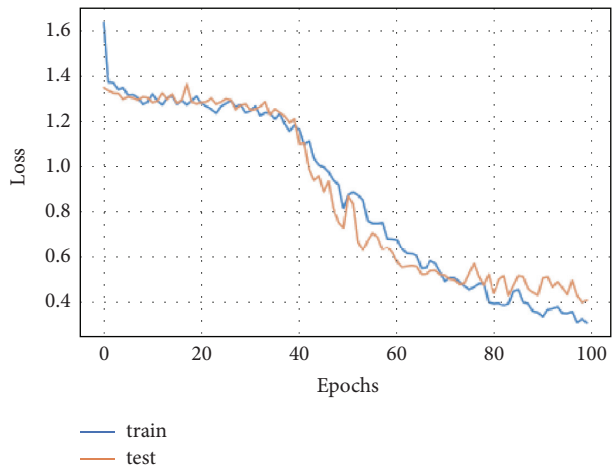


Figure 4: Model loss for 100 epochs.

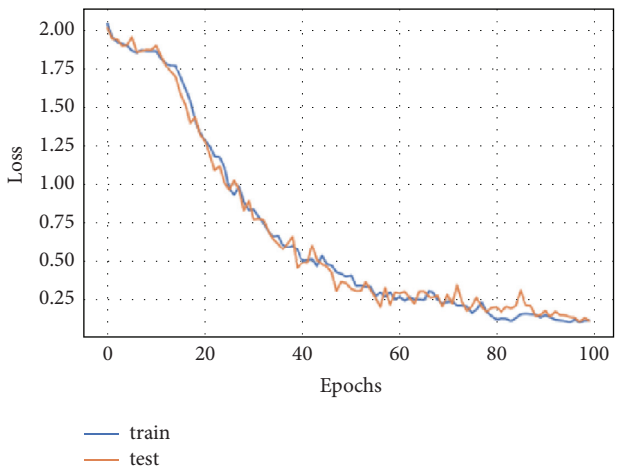


Figure 5: Model loss for 100 epochs with data duplication.

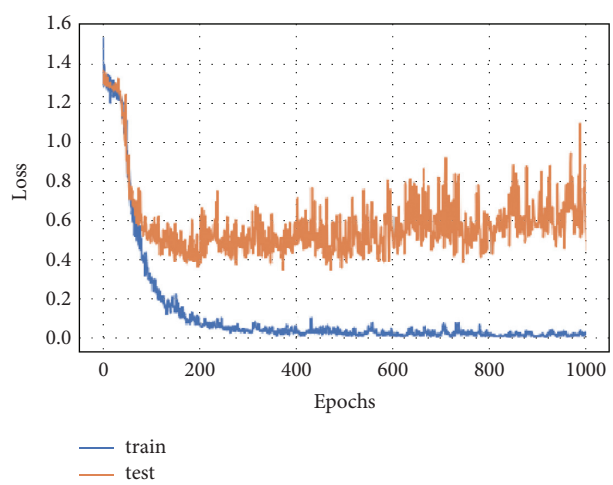


Figure 6: Model loss for 1000 epochs.

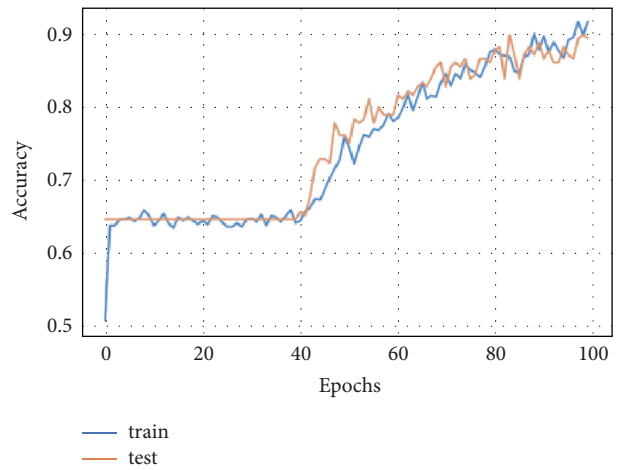


Figure 7: Model accuracy for 100 epochs.

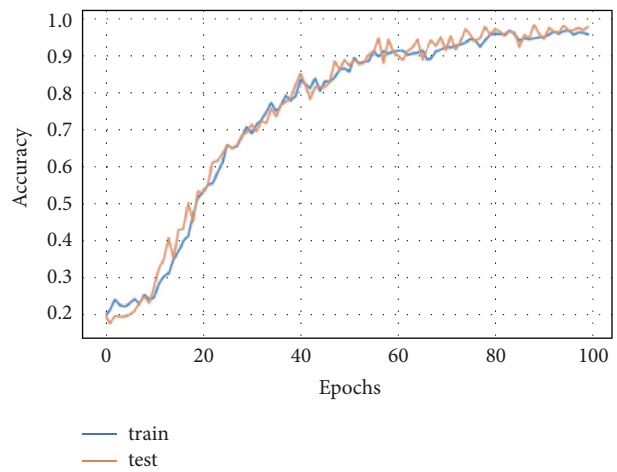


Figure 8: Model accuracy for 100 epochs with data duplication.

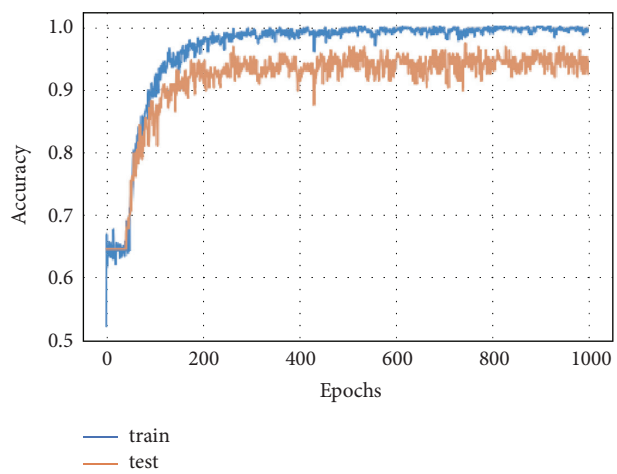


Figure 9: Model accuracy for 1000 epochs.

Table 5: Model accuracy.

Tests	Dataset	Accuracy (%)
Independent test set for 100 epochs	CK+	89.13
Independent test set for 100 epochs with data duplication	CK+	97
Independent test set for 1000 epochs	CK+	94
10 fold cross validation for 100 epochs	CK+	95.85
10 fold cross validation for 100 epochs with data duplication	CK+	99.79
10 fold cross validation for 1000 epochs	CK+	98.64

Table 6: Classification report.

Tests	Precision (%)	Recall (%)	F1 score (%)
Independent test set for 100 epochs	81.67	59.35	63.61
Independent test set for 100 epochs with data duplication	98.40	95.54	96.82
Independent test set for 1000 epochs	94.09	78.22	84.07
10 fold cross validation for 100 epochs	78.1	81	78.8
10 fold cross validation for 100 epochs with data duplication	99.10	98.40	98.40
10 fold cross validation for 1000 epochs	85.5	86.9	85.9

Table 7: Sensitivity and specificity.

Tests	Parameters	Results (%)
Independent test set for 100 epochs	Sensitivity	86.95
	Specifity	98.68
Independent test set for 100 epochs with data duplication	Sensitivity	97.55
	Specifity	99.69
Independent test set for 1000 epochs	Sensitivity	94.02
	Specifity	99.14
10 fold cross validation for 100 epochs	Sensitivity	83.28
	Specifity	97.63
10 fold cross validation for 100 epochs with data duplication	Sensitivity	99.17
	Specifity	99.88
10 fold cross validation for 1000 epochs	Sensitivity	94.53
	Specifity	99.22

help in better classification of multiclassification problems, as shown in Table 6.

5.2.5. Sensitivity and Specificity. The sensitivity and specificity of FDD-CNN are estimated with and without data duplication. The estimated metrics are presented in Table 7. The sensitivity evaluates the model's capability to predict true positives while specificity evaluates the model's capability to predict true negatives of every available category in the dataset, respectively.

5.2.6. Receiver Operating Characteristic (ROC) Curve. The ROC curve is used to evaluate the performance of multiclassification problems at different thresholds and, with ROC, the term area under the curve (AUC) is also frequently used. The ROC curve shows how well the given model is performing at each point. If AUC is higher, it shows that the given model is correctly predicting the available classes in the dataset under test; otherwise, it represents the inefficiency of the given model under the given scenario. The ROC curves of the proposed FD-CNN for 100 epochs without data duplication and with data duplication are shown in Figures 10 and 11, respectively. While in Figure 12, the ROC of the FD-CNN is shown for 1000 epochs without data duplication. The higher AUC shows the comparatively good performance of the proposed FD-CNN.

5.3. Comparison of Results. The comparison of the proposed FD-CNN with the state-of-the-art work intelligent FER frameworks using deep learning techniques is described in Table 8.

In Table 8, it is shown that the framework proposed by Georgescu et al. [25] achieved 87.76% and the learning model proposed by Nianyin et al. [42] achieved 89% accuracy by considering the dataset CK+ with duplication of data for 100 epochs. The proposed model FD-CNN with the same CK + dataset obtained an accuracy of 89.13% without duplication of data while 97% with duplication of data for 100 epochs, as did in [25, 42].

There are two major reasons for selecting [25, 42] to have a comparison. First one, the authors used the CK + dataset, and the other one is that deep neural network is used in both proposed frameworks. Specifically, in [42], the author used the DSAE model to train its dataset. DSAE is from a family of

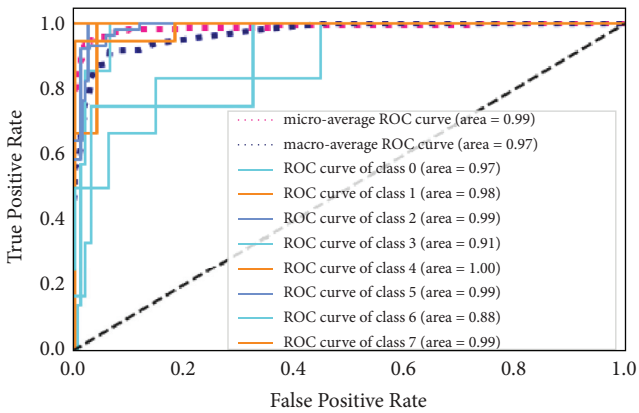


Figure 10: ROC curve for 100 epochs.

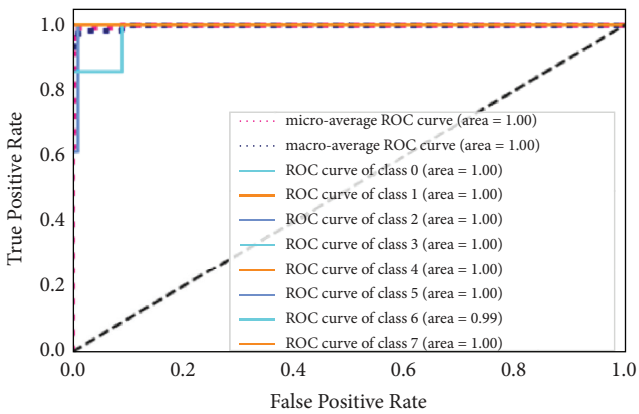


Figure 11: ROC curve for 100 epochs with data duplication.

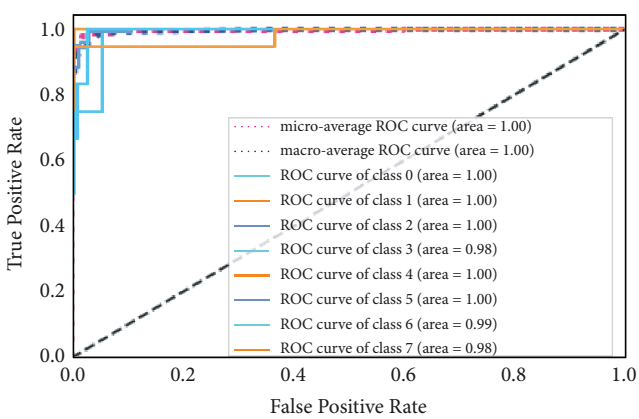


Figure 12: ROC curve for 1000 epochs.

Table 8: Comparison of Results

Dataset	Georgescu et al.	Nianyin et al.	FD-CNN
CK+	87.76%	89%	94%

deep neural networks. In this research work, CNN is used, which belongs to the deep neural network family and trains the model on the training dataset. To classify the images of the input dataset, [42] used the softmax classifier while in our proposed work, the softmax classifier is used to predict the true labels of images based on training. DSAE extracted the useful information from the input images, and the softmax classifier used this information for classifying the images among different categories. In the DSAE architecture, four layers are used in total. The first one is the input layer, the second and third are the hidden layers, and the fourth is the output layer. While in our proposed study, four convolutional layers, two hidden layers and one output layer are used. In [42], the features are extracted by using the histogram of oriented gradients (HOG) and these features are reduced with PCA. The model is then trained based on these features. While in our proposed FD-CNN, ReLu is used for the features that are to be extracted. Both studies are using backpropagation to fine-tune the model by minimizing the error rate. The total number of images of the CK + dataset with duplication used in [42] is 1635, while a total of 920 images of CK + without duplication and 1635 images with duplication are used in our proposed framework. For validation purposes, 10-fold cross-validation is used to evaluate the performance of the proposed model by measuring the parameters like sensitivity, specificity, precision, recall, and f1 measure.

6. Conclusions

In this study, an algorithm FD-CNN with four convolutional layers and two hidden layers is proposed for automatic FER to improve accuracy. Facial expression detection through facial images is used in many applications, including smart healthcare systems, criminology, holographic, security systems, education, robotics, entertainment, and stress detection. Compared to the previous work, where mostly six or seven expressions are used with different types of neural network. This research work used eight expressions for facial expression recognition. In this proposed method, the CK + dataset is used, with 123 different people, including males and females, with expressions such as anger, fear, happiness, sadness, disgust, contempt, neutral, and surprise. The proposed FD-CNN consists of preprocessing, feature extraction, and finally classification. In preprocessing, images are reshaped to 150×150 so that all images have an equal size. Furthermore, images are rotated randomly between 0 and 180 degrees, zoomed horizontally, and flipped vertically. After preprocessing, FD-CNN is used to extract the essential features. These extracted features are then used to classify facial expressions and send to the fully connected layer. Where through forward propagation, the error rate is estimated, and this error is then minimized by using backpropagation. In the end, the CNN classifier gives the label to these images based on the weighted score. The obtained accuracy of FDD-CNN for 100 epochs with data duplication is better than the state of the artwork as mentioned in the previous section.

The FD-CNN algorithm is also validated by using 10-fold cross-validation to validate its performance. This validation test includes precision, recall, and the f1 measure which is 94.09%, 78.22%, and 84.07%, respectively.

Furthermore, the calculated sensitivity and specificity are 94.02% and 99.14%, respectively.

Data Availability

The data are available from the following link: https://paperswithcode.com/dataset/ck.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] G. Muhammad, M. Alsulaiman, S. U. Amin, A. Ghoneim, and M. F. Alhamid, "A facial-expression monitoring system for improved healthcare in smart cities," *IEEE Access*, vol. 5, pp. 10871–10 881, 2017.

[2] B. A. Mercan, "Doing criminological research: affective states versus emotional reactions," *Theoretical Criminology*, vol. 24, no. 2, pp. 335–352, 2020.

[3] G. Giannakakis, M. Pediaditis, D. Manousos et al., "Stress and anxiety detection using facial cues from videos," *Biomedical Signal Processing and Control*, vol. 31, pp. 89–101, 2017.

[4] D. R. Faria, M. Vieira, F. C. Faria, and C. Premebida, "Affective facial expressions recognition for human-robot interaction," in *Proceedings of the 2017 26th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 805–810, IEEE, Lisbon, Portugal, August 2017.

[5] W. U. Khan, T. N. Nguyen, F. Jameel et al., "Learning-based resource allocation for backscatter-aided vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, no. 1–15, , 2021.

[6] A. Kharal, A. Mahmood, and K. Ullah, "Load forecasting of an educational institution using machine learning: the case of nust, islamabad," *Pakistan Journal of Science*, vol. 71, no. 4, p. 252, 2019.

[7] A. Mahmood, Y. Hong, M. K. Ehsan, and S. Mumtaz, "Optimal resource allocation and task segmentation in iot enabled mobile edge cloud," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 12, pp. 13294–13303, 2021.

[8] A. Mehrabian, "Communication without words," *Communication Theory*, vol. 6, pp. 193–200, 2008.

[9] P. Ekman, "An argument for basic emotions," *Cognition & Emotion*, vol. 6, no. 3-4, pp. 169–200, 1992.

[10] E. Friesen and P. Ekman, *Facial Action Coding System: A Technique For The Measurement Of Facial Movement*, Vol. 3, Consulting Psychologists Press, Palo Alto, CA, USA, 1978.

[11] M. Altamura, F. A. Padalino, E. Stella et al., "Facial emotion recognition in bipolar disorder and healthy aging," *The Journal of Nervous and Mental Disease*, vol. 204, no. 3, pp. 188–193, 2016.

[12] S. Happy and A. Routray, "Automatic facial expression recognition using features of salient facial patches," *IEEE transactions on Affective Computing*, vol. 6, no. 1, pp. 1–12, 2014.

[13] K. Masai, Y. Sugiura, M. Ogata, K. Kunze, M. Inami, and M. Sugimoto, "Facial expression recognition in daily life by embedded photo reflective sensors on smart eyewear," in *Proceedings of the 21st International Conference on Intelligent User Interfaces*, pp. 317–326, Sonoma, CA, USA, March 2016.

[14] N. Otberdout, A. Kacem, M. Daoudi, L. Ballihi, and S. Berretti, "Deep covariance descriptors for facial expression recognition," 2018, https://arxiv.org/abs/1805.03869.

[15] M. Sown, "A preliminary note on pattern recognition of facial emotional expression," in *Proceedings of the 4th International Joint Conferences on Pattern Recognition*, Kyoto, Japan, Novenber 1978.

[16] M. Pantic and L. J. M. Rothkrantz, "Facial action recognition for facial expression analysis from static face images," *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, vol. 34, no. 3, pp. 1449–1461, 2004.

[17] A. Johnston, B. B. Brown, and R. Elson, "Synchronous facial action binds dynamic facial features," *Scientific Reports*, vol. 11, no. 1, pp. 1–10, 2021.

[18] J. Chen, D. Chen, Y. Gong, M. Yu, K. Zhang, and L. Wang, "Facial expression recognition using geometric and appearance features," in *Proceedings of the 4th International Conference on Internet Multimedia Computing and Service*, pp. 29–33, Wuhan, China, September 2012.

[19] A. Uçar, Y. Demir, and C. Güzeliş, "A new facial expression recognition based on curvelet transform and online sequential extreme learning machine initialized with spherical clustering," *Neural Computing & Applications*, vol. 27, no. 1, pp. 131–142, 2016.

[20] L. Zhang and D. Tjondronegoro, "Facial expression recognition using facial movement features," *IEEE Transactions on Affective Computing*, vol. 2, no. 4, pp. 219–229, 2011.

[21] X. Zhao and S. Zhang, "Facial expression recognition based on local binary patterns and kernel discriminant isomap," *Sensors*, vol. 11, no. 10, pp. 9573–9588, 2011.

[22] T. Jabid, M. H. Kabir, and O. Chae, "Robust facial expression recognition based on local directional pattern," *ETRI Journal*, vol. 32, no. 5, pp. 784–794, 2010.

[23] M. Z. Uddin, M. M. Hassan, A. Almogren, A. Alamri, M. Alrubaian, and G. Fortino, "Facial expression recognition utilizing local direction-based robust features and deep belief network," *IEEE Access*, vol. 5, pp. 4525–4536, 2017.

[24] M. Jampour, V. Lepetit, T. Mauthner, and H. Bischof, "Pose-specific non-linear mappings in feature space towards multiview facial expression recognition," *Image and Vision Computing*, vol. 58, pp. 38–46, 2017.

[25] M.-I. Georgescu, R. T. Ionescu, and M. Popescu, "Local learning with deep and handcrafted features for facial expression recognition," *IEEE Access*, vol. 7, pp. 64827–64 836, 2019.

[26] K. Wang, X. Peng, J. Yang, D. Meng, and Y. Qiao, "Region attention networks for pose and occlusion robust facial expression recognition," *IEEE Transactions on Image Processing*, vol. 29, pp. 4057–4069, 2020.

[27] G. Bhullar and V. Mutneja, "Human face pose estimation based on feature extraction points," *International Journal of Computer Application*, vol. 142, no. 9, 2016.

[28] S. E. Gerard, J. Herrmann, Y. Xin et al., "Ct image segmentation for inflamed and fibrotic lungs using a multi-resolution convolutional neural network," *Scientific Reports*, vol. 11, no. 1, pp. 1–12, 2021.

[29] H. Kagaya, K. Aizawa, and M. Ogawa, "Food detection and recognition using convolutional neural network," in *Proceedings of the 22nd ACM International Conference on Multimedia*, pp. 1085–1088, Orlando, FL, USA, November 2014.

[30] M. Havaei, F. Dutil, C. Pal, H. Larochelle, and P.-M. Jodoin, "A convolutional neural network approach to brain tumor segmentation," in *BrainLes 2015,* Springer, Berlin, Germany, 2015.

[31] C.-C. Hsieh, M.-H. Hsih, M.-K. Jiang, Y.-M. Cheng, and E.-H. Liang, "Effective semantic features for facial expressions

recognition using svm," *Multimedia Tools and Applications*, vol. 75, no. 11, pp. 6663–6682, 2016.

[32] S. Jaiswal and G. Nandi, "Robust real-time emotion detection system using cnn architecture," *Neural Computing & Applications*, vol. 32, no. 15, pp. 11253–11262, 2020.

[33] Y. Huang, J. Yang, S. Liu, and J. Pan, "Combining facial expressions and electroencephalography to enhance emotion recognition," *Future Internet*, vol. 11, no. 5, p. 105, 2019.

[34] Z. Yu, G. Liu, Q. Liu, and J. Deng, "Spatio-temporal convolutional features with nested lstm for facial expression recognition," *Neurocomputing*, vol. 317, pp. 50–57, 2018.

[35] F. Zhang, Q. Mao, X. Shen, Y. Zhan, and M. Dong, "Spatially coherent feature learning for pose-invariant facial expression recognition," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 14, no. 1s, pp. 1–19, 2018.

[36] A. Qayyum, S. M. Anwar, M. Awais, and M. Majid, "Medical image retrieval using deep convolutional neural network," *Neurocomputing*, vol. 266, pp. 8–20, 2017.

[37] J. Ahmad, I. Mehmood, and S. W. Baik, "Efficient object-based surveillance image search using spatial pooling of convolutional features," *Journal of Visual Communication and Image Representation*, vol. 45, pp. 62–76, 2017.

[38] S. Bazrafkan, T. Nedelcu, P. Filipczuk, and P. Corcoran, "Deep learning for facial expression recognition: a step closer to a smartphone that knows your moods," in *Proceedings of the 2017 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 217–220, IEEE, Taiwan, China, July 2017.

[39] K. Low and U. Sheikh, "Human re-identification with global and local siamese convolution neural network," *Telkomnika*, vol. 15, no. 2, 2017.

[40] D. K. Jain, P. Shamsolmoali, and P. Sehdev, "Extended deep neural network for facial emotion recognition," *Pattern Recognition Letters*, vol. 120, pp. 69–74, 2019.

[41] M. Alsmadi, "Facial recognition under expression variations," *The International Arab Journal of Information Technology*, vol. 13, no. 1A, pp. 133–141, 2016.

[42] N. Zeng, H. Zhang, B. Song, W. Liu, Y. Li, and A. M. Dobaie, "Facial expression recognition via learning deep sparse autoencoders," *Neurocomputing*, vol. 273, pp. 643–649, 2018.

Hindawi
Journal of Healthcare Engineering
Volume 2022, Article ID 8950243, 13 pages
https://doi.org/10.1155/2022/8950243

Research Article

Tuberculosis Disease Diagnosis Based on an Optimized Machine Learning Model

Olfa Hrizi,[1] Karim Gasmi,[1] Ibtihel Ben Ltaifa,[2] Hamoud Alshammari,[3] Hanen Karamti,[4] Moez Krichen,[5] Lassaad Ben Ammar,[6] and Mahmood A. Mahmood[3]

[1]*Department of Computer Science, College of Arts and Sciences at Tabarjal, Jouf University, Jouf, Saudi Arabia*
[2]*STIH, Sorbonne Universite, Paris, France*
[3]*Department of Information Systems, College of Computer and Information Sciences, Jouf University, Jouf, Saudi Arabia*
[4]*Departement of Computer Sciences, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia*
[5]*Faculty of CSIT, Al-Baha University, Saudi Arabia & ReDCAD Laboratory, University of Sfax, Sfax, Tunisia*
[6]*College of Sciences and Humanities, Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia*

Correspondence should be addressed to Karim Gasmi; kgasmi@ju.edu.sa

Received 25 January 2022; Revised 21 February 2022; Accepted 1 March 2022; Published 21 March 2022

Academic Editor: Xingwang Li

Computer science plays an important role in modern dynamic health systems. Given the collaborative nature of the diagnostic process, computer technology provides important services to healthcare professionals and organizations, as well as to patients and their families, researchers, and decision-makers. Thus, any innovations that improve the diagnostic process while maintaining quality and safety are crucial to the development of the healthcare field. Many diseases can be tentatively diagnosed during their initial stages. In this study, all developed techniques were applied to tuberculosis (TB). Thus, we propose an optimized machine learning-based model that extracts optimal texture features from TB-related images and selects the hyper-parameters of the classifiers. Increasing the accuracy rate and minimizing the number of characteristics extracted are our goals. In other words, this is a multitask optimization issue. A genetic algorithm (GA) is used to choose the best features, which are then fed into a support vector machine (SVM) classifier. Using the ImageCLEF 2020 data set, we conducted experiments using the proposed approach and achieved significantly higher accuracy and better outcomes in comparison with the state-of-the-art works. The obtained experimental results highlight the efficiency of modified SVM classifier compared with other standard ones.

1. Introduction

Tuberculosis (TB) is a highly contagious disease mainly affecting the lungs (called "pulmonary TB"). When it affects other organs, it is called "extrapulmonary TB." It has rapidly spread all over the world and is currently considered one of the biggest threats to humanity. In 2015, the World Health Organization (WHO) estimated that TB caused over 1.8 million deaths worldwide [1, 2]. It is a clinical condition usually caused by a bacterium known as Mycobacterium tuberculosis [3]. Given that it affects multiple vital organs, it can be classified as a multisystem infectious disease. TB can also be classified as either "latent" or "active." In 2018, WHO estimated that 25% of the world's population suffers from latent TB [4]. From this organization, TB disease can be categorized into different types such as skeletal TB that spreads to bones from lungs. It is rare. Also, miliary TB affected the lungs and bone marrow but can spread to other parts of the body such as the brain or heart. Liver TB is considered from rare TB forms. It accounts for less than 1% of all TB infections. All of these types occurred after the appearance of some clinical symptoms such as coughing, fever, weight loss, and night sweats. So, it can be life-threatening if not properly treated. Therefore, early detection and diagnosis are the most effective prevention method. Automatic TB diagnosis and classification techniques have

been used in recent decades to improve the accuracy of disease recognition, thereby helping cardiologists make better decisions. Early techniques mainly used machine learning methods for automatic heartbeat classification, but required steps such as feature selection, feature extraction, and TB classification.

Machine learning (ML) is a branch of artificial intelligence that develops mathematical models using training data. Its objective is to give a precise diagnosis that makes decisions without being manually programmed for specific tasks. Mitchell gives an old, but still valid, definition in [5]. ML is of great interest to researchers owing to its ability to answer fundamental scientific and technical questions, as well as to improve the highly practical computer software produced and used in many applications [6–10], such as the medical field.

Several studies indicate that image processing techniques have been applied to diverse areas of research, such as security, engineering science, medical diagnosis [11], and film. Computer algorithms are used in image processing to enhance, restore, filter, classify, compress, segment, or threshold, enabling researchers to draw conclusions based on points of interest [12, 13]. Medical imaging—i.e., the process of visually portraying the inside of the human body [14]—can be used for medical treatment and clinical analysis. It is not just a basic method of detection, but can also provide diagnoses for various diseases.

This work concerns the medical imaging research field, focusing on TB. Diverse images have been developed for use with medical image technology, such as magnetic resonance imaging (MRI), computerized tomography (CT) scans, and X-rays. These technologies are necessary for accurate diagnostic imaging tests and are important in choosing the ideal treatment plan. All of them can be intensively analyzed and processed [12]. This technique is called image processing and is one of the branches of computer science [15]. It performs operations on images to extract data according to specific criteria and a well-ordered sequence of steps [16]. A representative technique for this process is segmentation, which divides a given image into multiple segments (sets of pixels, also known as super-pixels). Segmentation is typically used to identify objects or other relevant information in digital images. There are many different methods or algorithms for image segmentation [17].

Once image processing has been executed, the next step is the classification of the medical data. This stage was developed to overcome potential problems, with the objective of achieving promising results in the diagnosis of TB [18]. It is a system of categorizing all pixels within a digital image according to their characteristics, which are sorted into multiple sets of classes. Medical images are classified based on their extracted features. Input images may be classified as "normal" or "abnormal," and "benign" or "malignant." If each image is related to a unique one-class label or multi-label, the classification is single. If there are two class labels, the classification is binary; if there are more than two class labels, the classification is multiclass. Multiclass classification deals with a large number of labels [19], allowing each image to be tagged with more than one label. In this study, we use multiclass classification.

Feature extraction and representation is a crucial step in image processing, especially in the construction of any pattern classification. It consists of extracting the features that most accurately reflect the content of the images and assigning labels. During image classification, image features are extracted and numbered and then organized into classes. Some of the extracted features are irrelevant, redundant, or correlated, and sometimes, background noise leads to reduced efficiency and performance. Selecting the most meaningful features is a crucial step in classification. It is essential to remove insignificant features from the data set in order to guarantee more accurate diagnoses in medical applications. The performance of the classification model depends largely on the number of image features, which represent the input data set of the training model. However, there has been very little research on this problem. The extraction and selection of the most relevant features are still a challenge in the field of computer vision, particularly in image classification. Few works have classified TB using the K-nearest neighbor algorithm, a radial basis function (RBF) network, a multilayer perceptron (MLP) network, and a kernel regression to distinguish between bacilli and non-bacilli [20]. Support vector machine (SVM) is one of the most notable techniques [21, 22]. In this study, it is used in conjunction with a genetic algorithm to identify the smallest possible number of characteristics and to make the best possible distinctions between those features.

In this study, we propose an improved SVM based on an optimization algorithm for automatic TB classification. To improve the performance of the SVM classifier, we propose new image feature extraction and selection techniques that improve feature representation. In particular, we analyze the efficacy of fusing image feature extraction and selection techniques. The main contributions of this study are as follows:

- To quickly solve the quadratic optimization problem, particularly the problem of hyper-parameters, we invented an improved SVM classifier that attempts to classify the data set by finding its optimal parameters (C and γ).

- To improve the performance of the SVM classifier, we propose two different techniques—feature extraction and feature selection—to obtain the most relevant features. In feature extraction, potential distinguishing features were extracted using wavelet transform, which provides the most suitable scale to represent texture classification. The wavelet function offers a wide range of choices, whereas other options, such as Gabor filters, are less suitable due to their lack of orthogonality. The main purpose of the feature selection technique is to select a subset of input variables using cutout features with no predictive information while constructing a classification model. A genetic algorithm is used to identify the smallest possible collection of characteristics, allowing for the best possible discrimination between the retrieved features.

- To highlight the performance of the SVM classifier, this study offers a comparison between different machine learning algorithms based on accuracy rate.

The rest of this study is organized as follows: Section 2 reviews related works on TB detection techniques; our

materials and methods are described in Section 3; to conclude, Section 5 offers closing thoughts and a discussion of future work, while Section 4 describes the experimental assessment of our enhanced SVM classifier and compress the findings to those from other ML techniques.

2. Background

This section provides a brief description of support vector machines (SVMs) and then discusses related works in the fields of classification and disease diagnosis.

2.1. Support Vector Machine. Classification is a technique that trains with a suitable classifier to classify a given input. For our purposes, and given that our input data sets were not augmented, an SVM classifier was the better choice for tuberculosis (TB) classification. SVM is a powerful supervised classifier. It was first introduced to the statistical theory field in 1982 by Vapnick [23]; then, other studies demonstrated its effectiveness in various applications, such as medical diagnosis [24, 25]. Nowadays, SVM is used with a computational model, which increases its accuracy while decreasing its complexity, but also requires an improved SVM classifier. Its main goal is to construct optimal separating hyperplanes in higher dimensions, referred to as decision planes, as shown in Figure 1. Along with the hyperplanes, SVM transformed the original training data into multidimensional space for the purposes of classification. The middle line in Figure 1 represents the maximum margin hyperplane, which separates the two classes at the maximum distance from the closest data point [26].

The separating hyperplane can be linear or nonlinear.

2.2. Linear Separation. A hyperplane separates the input patterns in a linear type, presented by (2):

$$w^{\tau}X + b = 0, \tag{1}$$

where W^T is an adjustable weight vector and b is the bias term. For each training example x_i, we have the following:

$$Ef(x)\begin{cases} \geq 0 \text{for} y_i = +1, \\ \leq 0 \text{for} y_i = -1. \end{cases} \tag{2}$$

If $y = 1$, the input example is normal. If $y = -1$, the input example is abnormal. Suppose that for two hyperplanes H_1: $W'X_1 + b = 0$ and H_2: $W'X_2 + b = 0$, the smallest perpendicular distance to the data point from the hyperplane is calculated as $2/\|W\|$, and the best separation hyperplane is the one that maximizes the margin. The maximum margin hyperplane selection created by the SVM increases the accuracy of the classification and limits the number of misclassifications.

2.3. Nonlinear Separation. The linear SVM can be extended to a nonlinear classifier using a nonlinear operator $\varphi(.)$ to determine the input pattern x in a higher-dimensional space H. The nonlinear SVM classifier so obtained is defined as follows: it is possible to convert the linear SVM into a nonlinear classifier by employing the nonlinear operator varphi(.) to identify the input pattern x in a higher-dimensional space H, as shown in Figure 1. The following is the definition of the nonlinear SVM classifier that was obtained:

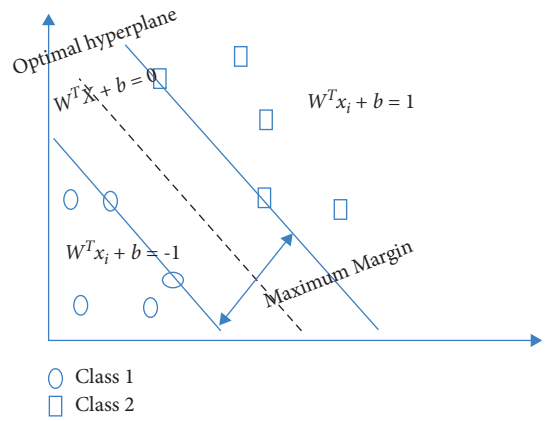


Figure 1: Separating hyperplane.

$$f(x) = W^T\varphi(X) + b. \tag{3}$$

The transformed data $\varphi(X)$ show linearity, but in terms of the original data $\vdots\ x \in R^n$, the classifier is nonlinear. To determine the parameters of decision function $f(x)$, it must follow the minimization criteria:

$$\text{Min}J(W, \varepsilon) = \frac{1}{2}\|W\|^2 + C\sum \varepsilon_I \text{Subjectto}: y_i\left(W^T\varphi(X) + b\right) \geq 1 - \varepsilon_i. \tag{4}$$

Following data preprocessing, a genetic algorithm (GA) selects a selection of features and extracts them using the spatial gray-level dependence method (SGLDM).

There are several common kernel functions, given: Linear $x_i \cdot x_i$ Polynomial of degree: d: $(x_i \cdot x_i + 1)^d$ Radial basis function (RBF) which is expressed as: $\exp((-\|x_i - x_j\|^2)/2\vartheta^2)$

2.4. Related Works. In this section, we rely on pure medical references to better explain TB, its location, its tendency to spread, its severity, and other data. This helps researchers to diagnose with greater precision and to use the right scientific and medical data in their studies.

In 2018, Kristen et al. [27] provided a systematic analysis of diverse bacterial strains that cause TB. It is also worth mentioning [5], in which the author concludes not only that TB is the multifaceted disease but also that there are many tests for its diagnosis. We also refer to studies in computer science, which use different techniques and algorithms to identify TB. We review below the main steps involved in a fully automatic TB detection system: digital image acquisition and preprocessing techniques, image segmentation methods, and feature extraction and classification.

2.5. Digital Image Acquisition and Preprocessing Techniques. In the literature, several studies have been presented for automatic TB detection, giving current researchers the ability to analyze the presence of TB bacteria automatically and quickly from a solid database [28]. An automatic TB detection system is used to automatically analyze the presence of TB bacteria quickly using different steps. The first step is image acquisition and preprocessing. Thus, before performing the image acquisition, an essential part of the automatic TB detection system called autofocusing has to be done [29] to save time and provide a better focused image. Several methods have been proposed for autofocusing [30, 31]. The performance of each method depends on different factors such as image characteristics, noise, and other specifics in the image.

Before processing the acquired images, a preprocessing step must be accomplished for the images to enhance their quality. Several preprocessing methods have been proposed to improve the contrast and brightness of an image. Susanto et al. [32] developed an approach to identify lung TB. Their study applied image preprocessing methods to make identification faster. In [33], the authors published the first systematic review in which diverse models were proposed for the prediction of TB treatment outcomes.

When the data set is not uniform and lacks fine textural features, the first image processing task is to obtain uniformity throughout the data set and improve the quality of the images. This technique also aims to reduce image backgrounds. Le in [34] used a small window to scan the lung region for TB classification. Additionally, image processing techniques such as image enhancement, segmentation, and feature extraction have been used in [35] for TB diagnosis.

To obtain input data, Poornimadevi et al. [36] employed X-ray images; in 2017, Antony and Banu [37] added filtering to the same input database. Other works used a computed tomography (CT) scan for ImageCLEF, which organized a challenge based on CT image analyses of TB patients in 2017 [38].

Another frequently cited technique is machine learning (ML), which is used to model training data. The authors in [39] recently evaluated ML models for their efficacy in estimating TB prognosis. Other researchers detailed the ML mechanism, citing the survey in [40] in the development of their concepts and applications. The authors in the last work defined ML as "the union of forces between statistics and computer sciences" and "the basis for artificial intelligence." In [41], the authors propose a feature selection model for brain tumor classification. In this study, we propose a model in two aspects, the first one for feature selection in the tuberculosis classification field and the second one for hyper-parameter selection. We aim to find the optimal SVM (C, γ) parameters to improve our classifier.

To study ML techniques such as logistic regression (LR) and linear discriminant analysis (LDA), we refer to [42] to identify causes, risk factors, and effective treatments.

Other recent researches such as [43, 44] adopted the ML algorithm for tuberculosis prediction.

2.6. Image Segmentation Methods. Image segmentation is one of the most crucial axes in medical image analysis. Its goal is to distinguish which objects caused a given disease within the tissue. Many TB-specific studies have proposed several segmentation algorithms such as thresholding methods, including K-means (KM) clustering, neural network-based approaches, and Bayesian segmentation [45, 46]. Several works proposed well-known thresholding algorithms to segment microscopic images, and other metaheuristic algorithms, such as KM, fuzzy c-means (FCM), fast marching (FM) thresholding, and the firefly algorithm (FA), were developed to solve bilevel microscopic image thresholding problems.

Other studies used neural network (NN)-based approaches to detect TB bacilli. Priya and Srinivasan [47] used digital TB images for image-level and object-level classifications based on the multilayer perceptron (MLP) NN. TB detection has also been implemented using deep learning-based processes [48]. Hwang et al. [49] developed a modified AlexNet and used transfer learning.

2.7. Feature Extraction Methods. After segmentation, some impurities and unwanted data may remain in the images, necessitating feature extraction. Feature extraction is the reduction in image feature values to obtain better results and greater speed during the classification process. Thus, the feature extraction process plays an important role in the design of a good classification model.

Techniques based on Fourier transform (FT) [12] and wavelet transform (WT) [14, 15] have been developed in this regard. WT performs time-frequency analysis, while FT merely examines frequency. This makes it a useful tool for pattern detection and time-space-frequency analysis.

2.8. Classification Methods. Small artifacts and undesired areas remain in the image even after segmentation has been performed. Different classification approaches, such as Bayesian classifiers, support vector machines (SVM), probabilistic neural networks (PNNs), and KNN classifiers, are used to extract true bacilli from these segments. The techniques presented in [50, 51] used traditional classifiers—namely the Bayesian, NN, and random forest (RF) classifiers—which seem to be inefficient in classifying overlapping bacilli. Hooda et al. [52] proposed three standard architectures (AlexNet, GoogleNet, and ResNet) to create a custom data set for TB classification. Chithra and Jagatheeswari [53] drew comparisons between different classifiers to verify that their fractional crow search-based support vector neural network was highly accurate and performed better than others.

On the other hand, we cannot deny that some studies resulted in poorly labeled approaches to classifying TB. With these approaches, the exact locations or outline boundaries cannot be provided to the system, but an image label is added to identify a given chest X-ray (CXR) as "normal" or "abnormal." These works are shown in [54, 55].

Other strategies were designed to detect multiple TB manifestations using a supervised approach, as with Jaeger

et al. [56], who described a technique to classify TB based on different shape and texture descriptors.

SVM is another important classifier that can be used to detect and count the number of TB bacteria. It is described in [57] as a state-of-the-art classifier in real-world pattern recognition applications. The current TB literature includes reviews comparing this classifier with more classical ones, where each comparison is based on some specific criterion, such as the outcome of each classifier.

In [21], SVM was compared with convolutional neural network (CNN) models. The experimental results showed that the best overall accuracy was 98.84%, obtained by an SVM-radial basis function (RBF) network model.

The authors in [58] proposed an automatic TB detection system that is based on the Gaussian fuzzy neural network (GFNN) classifier. The GFNN classifier combines the fuzzy classifier and the neural work with the Gaussian mixture model. It classifies the segments of the images into few bacilli, non-bacilli, and overlapping bacilli. The proposed GFNN model has achieved overall better performance in comparison with the various classical models, such as SVM, Bayesian regularization, Levenberg Marquardt, and fuzzy hyperbolic-based decision tree.

The authors in [59] proposed a hybrid classifier that combines the decision tree and the deep belief network along with the Gaussian model for infection-level identification in TB diagnosis. The classifier implements two-level classification techniques. At first-level classification, the images are categorized into three classes, such as few bacilli, non-bacilli, and overlapping bacilli. On the other hand, the second-level classification finds the number of bacilli by counting the bacilli and measuring the density ratio to determine the infection level. By showing a comparative analysis, the proposed Gaussian model has achieved overall better performance in comparison with that of existing conventional models, such as SVM, Bayesian regularization, Levenberg Marquardt, fuzzy hyperbolic-based decision tree, and GFNN.

One major challenge in TB screening is the development of a classification algorithm that guarantees higher accuracy, with a strong performance and sensitivity percentage to support doctors in making the right diagnosis.

3. Materials and Methods

In this section, we will introduce the materials we used and then describe the approaches we followed to achieve automatic tuberculosis (TB) classification.

3.1. Data Sets. Medical data collection is generally based on data volume, annotation, accuracy, and reusability. Each medical image can be defined by its data elements, metadata, and identifier. In this work, the data are provided as 3D computed tomography (CT) scans in compressed Neuroimaging Informatics Technology Initiative (NIfTI) files with the extension ".nii" from ImageCLEF campaign (https://www.imageclef.org/2020/medical/tuberculosis). After the files are decompressed, three sets of slices can be extracted, corresponding to the three dimensions of the 3D image XYZ (512 x 512 pixels). However, some experiments indicated that the slices of the Y dimension produced better results than that of X and Z dimensions; in fact, TB identification did not require every slice. Thus, we only have to keep those that are potentially informative. After selecting the Y dimension, the data set that we used contained 264 images. It is not a large database, but we chose it for the sake of a fully automated approach. In Figure 2, we give an example of a sample picture and the different types of TB-related findings.

4. Methodology

A whole system must be followed during medical image processing. It contains many functions and other iterative methods to perform the optimization algorithm. Our proposed methodology for TB detection is shown in Figure 3. The input is the CT scan image of a human lung. This image is then preprocessed to improve its quality, and the output is given to the feature extraction block, which generates the input for the classifier.

This system comprises four main stages:

(1) Input data preprocessing

(2) Feature extraction

(3) Feature selection

(4) SVM hyper-parameter selection.

Each step will be detailed throughout the rest of this section.

4.1. Preprocessing: Feature Extraction Using Wavelet Transform. There are several multi-resolution approaches to feature extraction, most prominently Fourier transform (FT) and wavelet transform (WT). We worked with WT, which is a general mathematical tool for analyzing complex data sets and signal processing. WT's capabilities include time-scale signal analysis, signal decomposition, and signal compression. A major drawback of the FT method is that it delivers the same frequency resolution throughout the window function's duration. Moreover, it does not capture the time-evolving impacts of frequencies in nonstationary signals, whereas WT functions do so by providing a hierarchy of scales, starting with the coarsest scale in either stationary or nonstationary signals, respectively. Consequently, WT was deemed the most suitable tool for feature extraction due to its capability of displaying a picture at a variety of various resolutions. WT is often produced from the mother wavelet, and it is one of the most widely used instances of frequency-domain analysis and computation.

For feature extraction, we used the spatial gray-level dependence method (SGLDM) developed by Haralick [60]. With this statistical technique, the geographical distribution of gray levels is estimated by computing the second-order conditional probability density g (i, j, d). An element at (i, j) of the SGLD matrix indicates the likelihood that two cells with differing resolutions, oriented away from the horizontal line, would have gray-level values I and j, respectively.

In Figure 4, WT image decomposition uses successive high-pass and low-pass filters.

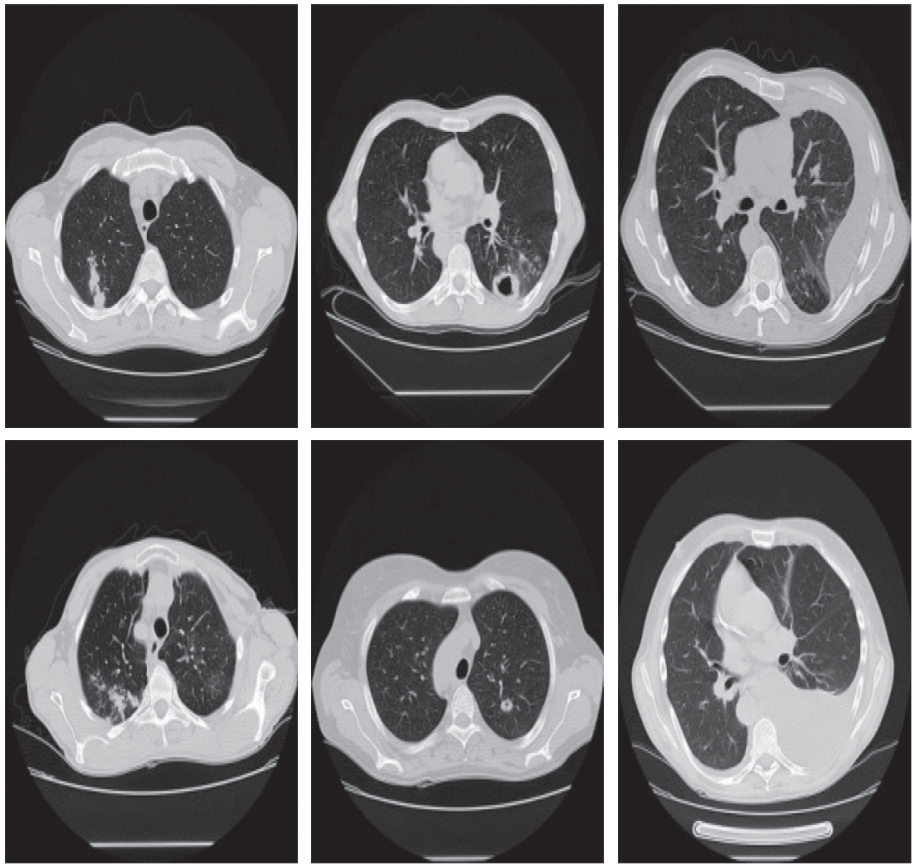


Figure 2: Slices of typical CT images with three types of TB-related findings (https://www.imageclef.org/2020/medical/tuberculosis).

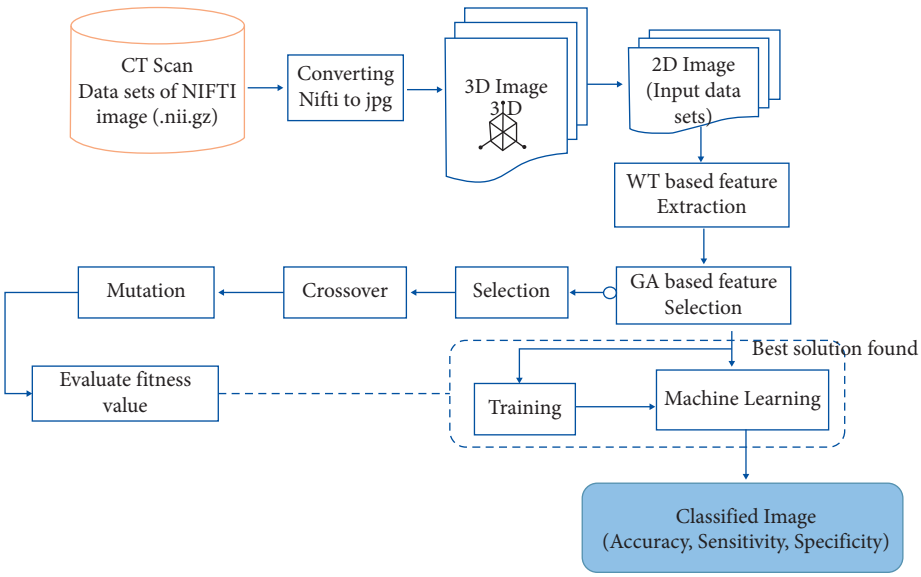


Figure 3: Illustration of the proposed approach.

22 features described in Table 1 were extracted by WT due to its multi-resolution capabilities.

4.2. Feature Selection via Genetic Algorithm. Those characteristics that are portrayed as actual numbers that exist in a high-dimensional space are not always meaningful or significant. Some of them are irrelevant, redundant, correlated, and occasionally noisy, making the learning models more likely to be overfit, complicated, and difficult to understand. As a result of these characteristics, data mining applications have low efficiency and poor performance (e.g.,

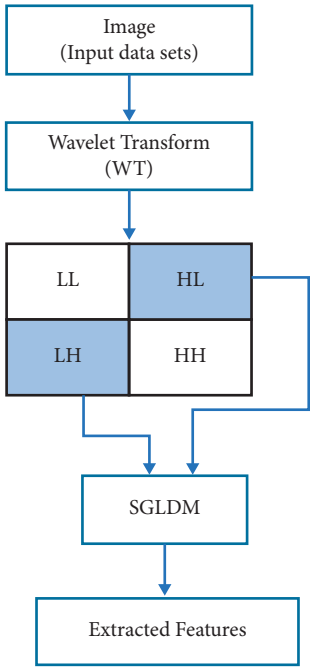


Figure 4: Block diagram of feature extraction by applying WT and SGLDM.

Table 1: Extracted features.

Features (mean/range)	
Frequency domain (32,33, 34)	Spatial domain (32,35)
Angular second moment	Correlation
Contrast	Variance
Inverse difference moment	Sum variance
Sum average	Difference variance
Entropy, sum entropy, difference entropy	Information measure of correlation I
Cluster prominence	Information measure of correlation II
Cluster shade	Maximal correlation coefficient
Dissimilarity	Correlation matrix
Homogeneity and inverse difference normalized	Maximum probability
Energy	

classification). The reduction in data dimensionality and the reduction in computational complexity are both required to choose robust features from unlabelled data to solve these difficulties.

Feature selection aims to select the most important one of the original features to avoid overfitting. To accomplish this task, we used a new technique based on a genetic algorithm (GA) for unsupervised feature selection.

GAs are stochastic search techniques, based on natural genetics, which provide robust search capabilities in complex spaces. A GA is an iterative process that solves an optimization problem [44]. Each solution is obtained by means of an encoding/decoding mechanism, which requires us to represent the solution as a chromosome. This is repeated conversely. To indicate the lack or existence of a feature, it is represented by zero or one at position i. GAs begin with a randomly generated population of chromosomes. A fitness function measures the solution's quality and efficacy. So, the fitness function in (1) treats the chromosome as input and outputs its fitness value. The next phase is to choose the fittest individuals to be future parents. Individuals can be used to create new populations by applying reproductive operators such as crossover and mutation.

$$\text{fitness} = W_{\text{acc}} \times \text{Accuracy} + W_{ft} \times \frac{1}{N}, \tag{5}$$

where W_acc represents accuracy weight, W_ft represents feature weight, and N represents the number of features selected.

CT scans for tuberculosis (TB) can benefit from applying GAs to guarantee that the best feature sets are selected for analysis. Competition among feature transformation matrices is maintained. For each matrix in this population, transformed patterns are computed by multiplying the input patterns by the matrix to obtain a collection of chosen features, as illustrated in Figure 5. These features are then given to a classifier.

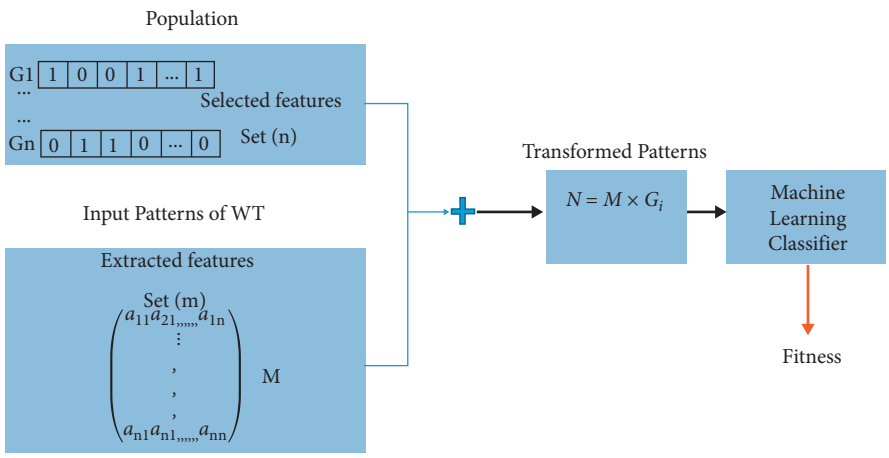


Figure 5: Genetic algorithm.

The GA system seeks to discover m optimal features from n extracted features to improve classifier performance. Finally, the GA's major purpose is to reduce the dimensionality of modified patterns while increasing classification accuracy.

4.3. SVM Parameter Selection. A large number of kernel functions are used to aid the support vector machine (SVM) in its pursuit of the best solution. The polynomial, sigmoid, and radial basis function (RBF) kernels are often employed. Unlike linear kernel functions, RBF is the most commonly used kernel function because it can effectively categorize multidimensional data. To get the best possible results, we used an RBF kernel function in our SVM. C and γ, two of the most important RBF parameters used with SVMs, had to be adjusted accordingly. In the example above, the parameter C represents the cost of the penalty. Because the value of this parameter has an influence on the partitioning results in the feature space, the value of this parameter has a substantially bigger impact on classification outcomes than the value of the other parameters.

The best classification accuracy rate may be determined by selecting suitable values for the upper and lower limits (search interval), as well as for the jumping interval, during the search process. In addition to the parameters C and γ, additional variables, such as the quality of the features data set, may have an impact on the classification accuracy rate, such as the number of false positives. Examples of this include the correlations between characteristics, which have an impact on the categorization outcome.

It is necessary to tweak C and γ when using the RBF kernel. It has been established that improper selection of the two parameters might result in over- or underfitting in the model. The suggested GA-based technique is intended to optimize C and γ for the SVM. This study employs the RBF kernel function for the SVM classifier to accomplish our suggested technique. Classification accuracy and the quantity of chosen characteristics were utilized to build a fitness function.

The proposed GA-SVM approach is described in greater detail below:

Step 1. A binary string representation of the parameters and chromosomes that reflects the SVM parameters is created.

Step 2. The initial population of chromosomes is randomly generated, and then, the population is initialized.

Step 3. The parameters that were chosen (C and γ) are located.

Step 4. To get the trained optimal SVM, you will need to run your data through it with the new parameters. This model predicts the test sets.

Step 5. Fitness is assessed. To compute the k-fold cross-validation accuracy, the optimum chromosomes and optimal pair (C, gamma) are entered into an SVM classifier for each chromosome.

Step 6. The maximum number of generations N must be reached before the end of the game, or the fitness value of the preceding M generations must be lower than the current generation. There is no further iteration is possible if both conditions are satisfied.

5. Results and Discussion

All proposed techniques were implemented on computed tomography (CT) scans of tuberculosis (TB) patients. The input data set was divided into the training and testing sets. After the data set was collected, feature extraction was applied using the spatial gray-level dependence method (SGLDM) technique. In this section, the findings obtained through experimental studies are presented in two sections: (1) the findings obtained using hyper-parameter selection and (2) the findings obtained using feature selection.

Python was used to develop the model, which was tested using RTX 2060 Graphics Card and 16 GB of RAM.

5.1. Data Set and Evaluation Metrics. We used the ImageCLEF 2020 (https://www.imageclef.org/2020/medical/tuberculosis) data set, which is freely available on the

Internet, to evaluate our enhanced machine learning model. In this study, we used multi-label classification. Three labels were assigned to each lung: "lung affected," "presence of pleurisy," and "presence of caverns." The "left lung affected" and "right lung affected" labels marked the presence of any kind of tuberculosis (TB)-associated damage in the left and right lungs, respectively. Tables 2 and 3 detail the distribution of patients within each label.

In this study, we were most interested in accuracy, which is a metric for evaluating the overall efficacy of the classifier. The accuracy metric measures the probability that a diagnostic test was correctly performed (i.e., the proportion of correctly classified images). It can be evaluated using the following formula:

$$|\text{Acc} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})}, \quad (6)$$

where TP indicates that the model correctly predicted the positive class. FP indicates that the model incorrectly predicted the positive class. FN indicates that the model incorrectly predicted the negative class. TN indicates that the model correctly predicted the negative class. All of these values are defined in Table 4.

Using the model described in this research, we conducted a series of experiments. Comparing the outcomes of various machine learning methods is also part of our research, for that we use Sklearn (https://scikit-learn.org/stable/) as Python package. scikit-learn (also known as Sklearn) is the most useable and stable machine learning toolkit available for Python. This package contains a range of efficient tools for machine learning and statistical modeling, such as classification and regression, clustering, and dimensionality reduction, all of which are accessible through a consistent interface in the Python programming language.

5.2. SVM Hyper-Parameter Selection. To improve the performance of the support vector machine (SVM), it is critical to first decide which parameters should be employed. To do so, we used an adaptive genetic algorithm (GA) to discover the ideal parameters. The primary purpose of the GA was to select the optimal parameters from a large number of alternative values to ensure the maximum possible accuracy throughout the training and testing stages. The parameters of the GA are presented in Table 5.

As per the discussion of kernel functions at the end of Section 3, the radial basis function (RBF) kernel is the most useful in terms of defining the optimal values of constants γ and C. Note that γ is the width of the kernel function and C is the error/trade-off parameter, which adjusts the significance of the separation error in the creation of the separation surface. The methodology proposed in this work was evaluated based on its accuracy. Table 6 shows the variation in classification accuracy scores, which ranged from 0.64 to 0.97. All values were assessed using the validation data set.

Here, we present some of the many suggestions resulting from variations in the parameters selected by the GA to show the limits of the accuracy rates for each class. We note that the mean accuracy obtained by the SVM classifier reached its maximum values, as compared to the parameters individually chosen by the GA. As noted in the previous section, the SVM classifier was approved for this study due to the moderate size of the data set.

Table 2: Class distribution in tuberculosis data set.

Label	Number of occurrences	Test
Left lung affected	211	75
Right lung affected	233	99
Caverns left	66	3
Caverns right	79	5
Pleurisy left	7	28
Pleurisy right	14	46

Table 3: The percentage of train and test data for each label.

	Left lung affected	Right lung affected	Left lung pleurisy	Right lung pleurisy	Left lung caverns	Right lung caverns
Train	211 (75%)	233 (82%)	7 (2%)	14 (4%)	66 (23%)	79 (28%)
Test	75 (63%)	99 (83%)	3 (3%)	5 (4%)	28 (23%)	46 (38%)

Table 4: Truth table.

	Outcomes	Disease	
Test	+	TP	FP
	-	FN	TN

Table 5: GA parameters.

GA property	Value/method
Size of generation	100
Initial population size	30
Selection method	Tournament
Number of crossover points	1
Crossover probability	0.9
Mutation method	Uniform mutation
Mutation probability	0.05

5.3. Comparison between SVM and Best-Known Machine Learning Models. In this section, we compare performance algorithms based on accuracy rate. The classifiers are as follows:

- "KNN," K-neighbor classifier [61].
- "CART," decision tree classifier [62].
- "NB," Gaussian NB [63].
- "LDA," linear discriminant analysis [64].
- "RF," random forest classifier [65].

In Table 7, we compared our SVM classifier with the supervised machine learning classifiers KNN, CART, NB, LDA, and RF. The experimental results show that our SVM classifier was more accurate than the other classification

Table 6: Sample of binary classification results based on an adaptive genetic algorithm.

	Accuracy					
	Left lung affected	Right lung affected	Caverns left	Caverns right	Pleurisy left	Pleurisy right
$C = 1$, $\gamma = 35$	0.76	0.82	0.78	0.74	0.97	0.95
$C = 20$, $\gamma = 35$	0.68	0.75	0.64	0.64	0.95	0.93
$C = 8$, $\gamma = 35$	0.69	0.74	0.69	0.65	0.95	0.94
$C = 1.5$, $\gamma = 20$	0.75	0.81	0.76	0.73	0.97	0.95
$C = 1.5$, $\gamma = 100$	0.73	0.81	0.78	0.71	0.97	0.95

Table 7: Comparison between performance algorithms based on the rate of accuracy.

	SVM	KNN	CART	NB	LDA	RF
Left lung affected	0.77	0.72	0.63	0.46	0.72	0.70
Right lung affected	0.82	0.80	0.69	0.74	0.81	0.81
Caverns left	0.78	0.76	0.68	0.72	0.75	0.75
Caverns right	0.77	0.73	0.51	0.42	0.76	0.68
Pleurisy left	0.97	0.97	0.95	0.82	0.95	0.97
Pleurisy right	0.95	0.95	0.93	0.89	0.93	0.95
Mean accuracy	0.84	0.82	0.73	0.67	0.82	0.81

Table 8: Sample of selected feature results for the left and right lung affected category.

Range of selected features	Left lung	Left lung
[1 : 20]	0.75	0.79
[3 : 12]	0.78	0.83
[3 : 14]	0.77	0.82
[3 : 20]	0.77	0.81
[3 : 10]	0.75	0.8

algorithms, while KNN and LDA performed better than CART and NB.

5.4. Feature Selection. Table 8 gives the range of selected features, as well as the accuracy score obtained in each range. Using the selection method, the optimal features were selected based on a GA and then used as input for the SVM classifier. The highest accuracy was obtained with the [3 : 12] range. By filtering the range of selected features, we minimized the number of extracted features, thereby accelerating the process of training.

As noted in the previous section, the SVM classifier was approved for this study due to the moderate size of the data set. Our modified SVM classifier, based on optimal parameters and feature-based selection methods, strongly improves classification accuracy. The optimal features used as input for our classifier were selected using a GA. The experimental results show that the GA was able to minimize the dimensionality of the transformed patterns while maximizing classification accuracy. It is very clear that our SVM classifier has achieved satisfactory results and attained a high classification accuracy rate. We can therefore note that our SVM classifier attempts to classify data sets by finding an optimal hyperplane and thus solves a quadratic optimization problem. In conclusion, our optimal SVM model significantly outperformed other models in classifying TB. The combination of optimal parameters and feature-based selection methods improved the performance of the GA in extracting robust and significant features. In general, the performance of machine learning algorithms is heavily dependent on the set of features to which they are applied.

6. Conclusion

In this study, we dealt with the problem of tuberculosis (TB) disease classification. Our main conclusions are as follows: the methodology discussed in this study highlights several techniques used in the field of medical image processing. Wavelet transform was used in conjunction with the spatial gray-level dependence method to extract features from the data set. These were then selected using an optimization genetic algorithm (GA) and were used as input for the support vector machine (SVM) classifier. To improve the performance of the SVM classifier, we used two new techniques based on our GA. For the first technique, we used an adaptive GA to determine the optimal parameters from a range of values in order to guarantee the highest possible accuracy during the training and testing phases. The second technique was used to select a small number of original features (i.e., the most one's features) to avoid overfitting and reduce the dimensional of the data. Finally, our experimental results show that our modified SVM classifier was more accurate than other classification algorithms in classifying TB. This study proves that the quality of extracted features has a direct impact on the effectiveness of image classification. Additionally, the proposed classification model can help doctors automatically diagnose TB because it possesses all of the qualities described in this study: accurate, robust, and easy to control. In future works, several improvements can be made. In particular, we are planning to address hybrid deep learning-based TB detection for improving the obtained results.

Data Availability

The Tuberculosis "TB 2020" data set used to support the findings of this study were supplied by the ImageCLEF campaign under license and so cannot be made freely available. Requests for access to these data should be made to https://www.imageclef.org/2020/medical/tuberculosis/; all the information necessary to have the database and the access authorization are mentioned in the following link: https://www.imageclef.org/2020/medical/tuberculosis/. The necessary codes for the machine learning models are

mentioned in the following link: https://scikit-learn.org/stable/

Disclosure

The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

The authors extend their appreciation to the Deanship of Scientific Research at Jouf University for funding this work through research grant no. DSR-2021-02–0109 and the Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R192), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

References

[1] World Health Organization, *Global Tuberculosis Report 2015*, World Health Organization, Cairo, Egypt, 20th edition, 2015.

[2] A. Yakobi, J. Z. Porterfield, J. Toman et al., "Hiv, tuberculosis, and otogenic intracranial sepsis: a devastating disease with a subtle presentation," *Otology & Neurotology*, vol. 40, no. 7, pp. e704–e712, 2019.

[3] World Health Organization, "Tuberculosis," 2021, http://www.who.int/es/news-room/fact-sheets/detail/tuberculosis.

[4] D. M. Iseman, "Tuberculosis: history," 2013, https://www.nationaljewish.org/conditions/tuberculosis-tb/history.

[5] M. Tom, *Mitchell. Machine Learning*, McGraw-Hill, New York, NY, USA, 1997.

[6] Q. Abu Al-Haija, M. Krichen, and W. Abu Elhaija, "Machine-learning-based darknet traffic detection system for iot applications," *Electronics*, vol. 11, no. 4, p. 556, 2022.

[7] S. Mian Qaisar, N. Alyamani, A. Waqar, and M. Krichen, "Machine learning with adaptive rate processing for power quality disturbances identification," *SN Computer Science*, vol. 3, no. 1, pp. 1–6, 2022.

[8] S. Qaisar, A. Mihoub, M. Krichen, and H. Nisar, "Multirate processing with selective subbands and machine learning for efficient arrhythmia classification," *Sensors*, vol. 21, no. 4, p. 1511, 2021.

[9] S. Srinivasan, V. Ravi, V Sowmya, M. Krichen, D. B. Noureddine, and K. P. Soman, "Deep convolutional neural network based image spam classification," in *Proceedings of the 2020 6th conference on data science and machine learning applications (CDMA)*, pp. 112–117, IEEE, Riyadh, Saudi Arabia, March 2020.

[10] A. Mihoub, H. Snoun, M. Krichen, R. Bel Hadj Salah, and M. Kahia, "Predicting covid-19 spread level using socio-economic indicators and machine learning techniques," in *Proceedings of the 2020 First International Conference of Smart Systems and Emerging Technologies (SMARTTECH)*, pp. 128–133, IEEE, Riyadh, Saudi Arabia, November 2020.

[11] K. Gasmi, I. B. Ltaifa, G. Lejeune, H. Alshammari, L. B. Ammar, and M. A. Mahmood, "Optimal deep neural network-based model for answering visual medical question," *Cybernetics & Systems*, pp. 1–22, 2021.

[12] N. Walia, H. Singh, S. Kumar Tiwari, and A. Sharma, "A decision support system for tuberculosis diagnosability," *International Journal of Soft Computing*, vol. 6, no. 3, pp. 1–13, 2015.

[13] M. Venu Madhavan, D. Ngoc Hoang Thanh, A. Khamparia, S. Pande, R. Malik, and D. Gupta, "Recognition and classification of pomegranate leaves diseases by image processing and machine learning techniques," *Computers, Materials & Continua*, vol. 66, no. 3, pp. 2939–2955, 2021.

[14] Y. Zhou, *Medical Imaging: Principles and Applications*, IntechOpen Book Series, London, UK, 2019.

[15] M. P. A. Kamble, M. V. V. Anagire, and M. S. N. Chamtagoudar, "Cxr tuberculosis detection using matlab image processing," *International Research Journal of Engineering and Technology*, vol. 3, 2016.

[16] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice-Hall, Hoboken, NJ, USA, 2008.

[17] L. Wang, R. Chen, S. Wang, N. Zeng, X. Huang, and C. Liu, "Nested dilation network (ndn) for multi-task medical image segmentation," *IEEE Access*, vol. 7, pp. 44676–44685, 2019.

[18] A. Krizhevsky, I. Sutskever, and G. E Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds., Vol. 25, Curran Associates, Inc., New York, NY, USA, 2012.

[19] R. Alazaidah and F. Kabir Ahmad, "Trending challenges in multi label classification," *International Journal of Advanced Computer Science and Applications*, vol. 7, p. 10, 2016.

[20] M. K. Osman, M. Y. Mashor, and H. Jaafar, "Performance comparison of extreme learning machine algorithms for mycobacterium tuberculosis detection in tissue sections," *Journal of Medical Imaging and Health Informatics*, vol. 2, no. 3, pp. 307–312, 2012.

[21] H. Hasan, H. Z. M. Shafri, M. Habshi, and M. Habshi, "A comparison between support vector machine (SVM) and convolutional neural network (CNN) models for hyperspectral image classification," in *Proceedings of the IOP Conference Series: Earth and Environmental Science*, vol. 357, no. 1, Article ID 012035, Bandung, Indonesia, November 2019.

[22] F. Rubio, J. Martínez-Gómez, M. Julia Flores, and J. M. Puerta, "Comparison between bayesian network classifiers and svms for semantic localization," *Expert Systems with Applications*, vol. 64, pp. 434–443, 2016.

[23] V. Vapnik, *Estimation of Dependences Based on Empirical Data*, Springer Science & Business Media, Heidelberg, Germany, 2006.

[24] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Machine Learning*, vol. 46, pp. 389–422, 2002.

[25] J. Zhang and Y. Liu, "Cervical cancer detection using svm based feature screening," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, vol. 3217, pp. 873–880, Saint-Malo, France, September 2004.

[26] D. Morariu, N. V. Lucian, and V. Tresp, "Feature selection methods for an improved svm classifier," *International Journal of Computer and Information Engineering*, vol. 2, 2006.

[27] K. E. Wiens, L. P. Woyczynski, J. R. Ledesma et al., "Global variation in bacterial strains that cause tuberculosis disease: a

systematic review and meta-analysis," *BMC Medicine*, vol. 16, p. 196, 2018.

[28] R. O. Panicker, B. Soman, G. Saini, and J. Rajan, "A review of automatic methods based on image processing techniques for tuberculosis detection from microscopic sputum smear images," *Journal of Medical Systems*, vol. 40, no. 1, pp. 17–13, 2016.

[29] M. J Russell and T. S Douglas, "Evaluation of autofocus algorithms for tuberculosis microscopy," in *Proceedings of the 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3489–3492, IEEE, Lyon, France, August 2007.

[30] C. F. Fernandes Costa Filho, M. G. Fernandes Costa, and A. Kimura Junior, "Autofocus functions for tuberculosis diagnosis with conventional sputum smear microscopy," *Current Microscopy Contributions to Advances in Science and Technology*, pp. 13–20, 2012.

[31] S. Pertuz, D. Puig, and M. A. Garcia, "Analysis of focus measure operators for shape-from-focus," *Pattern Recognition*, vol. 46, no. 5, pp. 1415–1432, 2013.

[32] R. N. Rohmah, A. Susanto, I. Soesanti, and M. Tjokronagoro, "Computer aided diagnosis for lung tuberculosis identification based on thoracic x-ray," in *Proceedings of the 2013 International Conference on Information Technology and Electrical Engineering (ICITEE)*, pp. 73–78, IEEE, Yogyakarta, Indonesia, October 2013.

[33] L. S. Peetluk, F. Moreira Ridolfi, P. F. Rebeiro, D. Liu, V. C. Rolla, and T. R. Sterling, "Systematic review of prediction models for pulmonary tuberculosis treatment outcomes in adults," *BMJ Open*, vol. 11, 2021.

[34] L. Kim, "Automated detection of early lung cancer and tuberculosis based on x-ray image analysis," in *Proceedings of the 6th WSEAS International Conference on Signal, Speech and Image Processing, SSIP'06*, pp. 110–115, Lisbon, Portugal, September 2006.

[35] M. ullah, M. Bari, A. Ahmed, and S. Naveed, "Lungs cancer detection using digital image processing techniques: a review," *Mehran University Research Journal of Engineering and Technology*, vol. 38, no. 2, pp. 351–360, 2019.

[36] C. S. Poornimadevi and H. Sulochana, "Automatic detection of pulmonary tuberculosis using image processing techniques," in *Proceedings of the 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, pp. 798–802, IEEE, Chennai, India, March 2016.

[37] B. Antony, "Lung tuberculosis detection using x-ray images," *International Journal of Applied Engineering Research*, vol. 12, 2018.

[38] Y. Dicente Cid, V. Liauchuk, V. Kovalev, and H. Müller, *Overview of Imageclef 2017 Tuberculosis Task – Predicting Tuberculosis Type and Drug Resistances*, CLEF, Dublin, Ireland, 2017.

[39] M. H. L. Ferreira Da Silva Barros, G. Oliveira Alves, L. M. Florêncio Souza et al., "Benchmarking machine learning models to assist in the prognosis of tuberculosis," *Informatics*, vol. 8, no. 2, 2021.

[40] R. Behera and K. Das, "A survey on machine learning: concept, algorithms and applications," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 2, no. 2, 2017.

[41] K. Ahmed, K. Gasmi, M. Anouar Ben Messaoud, N. Benamrane, and A. Mohamed, "A hybrid approach for automatic classification of brain mri using genetic algorithm and support vector machine," *Leonardo Journal of Sciences*, vol. 17, no. 1, 2010.

[42] A. Callahan and N. H. Shah, "Machine learning in healthcare," in *Key Advances in Clinical Informatics*, Academic Press, pp. 279–291, Cambridge, MA, USA, 2017.

[43] O. Faruk, E. Ahmed, S. Ahmed et al., "A novel and robust approach to detect tuberculosis using transfer learning," *Journal of healthcare engineering*, vol. 2021, Article ID 1002799, 10 pages, 2021.

[44] M. H. Ali, D. M. Khan, K. Jamal, Z. Ahmad, S. Manzoor, and Z. Khan, "Prediction of multidrug-resistant tuberculosis using machine learning algorithms in swat, Pakistan," *Journal of healthcare engineering*, vol. 2021, Article ID 2567080, 11 pages, 2021.

[45] R. A. A. Raof, M. Yusoff Mashor, and S. S. Md Noor, "Segmentation of Tb Bacilli in Ziehl-Neelsen Sputum Slide Images Using K-Means Clustering Technique," 2018.

[46] S. Ayas, H. Dogan, E. Gedikli, and M. Ekinci, "A novel approach for bi-level segmentation of tuberculosis bacilli based on meta-heuristic algorithms," *Advances in Electrical and Computer Engineering*, vol. 18, no. 1, 2018.

[47] E. Priya and S. Srinivasan, "Separation of overlapping bacilli in microscopic digital tb images," *Biocybernetics and Biomedical Engineering*, vol. 35, no. 2, pp. 87–99, 2015.

[48] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–22, 2021.

[49] S. Hwang, H.-E. Kim, J. Jeong, and H.-J. Kim, "A novel approach for tuberculosis screening based on deep convolutional neural networks," in *Proceedings of the Medical Imaging 2016: Computer-Aided Diagnosis, of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 9785, San Diego, CA, USA, March 2016.

[50] E. Priya and S. Srinivasan, "Automated object and image level classification of tb images using support vector neural network classifier," *Biocybernetics and Biomedical Engineering*, vol. 36, p. 7, 2016.

[51] J. Singh, A. Tripathy, P. Garg, and A. Kumar, "Lung tuberculosis detection using anti-aliased convolutional networks," in *Proceedings of the International Conference on Smart Sustainable Intelligent Computing and Applications under ICITETM2020*, vol. 173, pp. 281–290, Delhi, India, February 2020.

[52] R. Hooda, A. Mittal, and S. Sofat, "Automated tb classification using ensemble of deep architectures," *Multimedia Tools and Applications*, vol. 78, p. 11, 2019.

[53] R. S Chithra and P. Jagatheeswari, "Fractional crow search-based support vector neural network for patient classification and severity analysis of tuberculosis," *IET Image Processing*, vol. 13, p. 10, 2018.

[54] Y. Arzhaeva, L. Hogeweg, P. Jong, M. Viergever, and B. Ginneken, "Global and local multi-valued dissimilarity-based classification: application to computer-aided detection of tuberculosis," *Med Image Comput Comput Assist Interv*, vol. 12, pp. 724–731, 2009.

[55] J. Melendez, B. Van Ginneken, P. Maduskar, R. H. H. M. Philipsen, H. Ayles, and C. I. Sanchez, "On combining multiple-instance learning and active learning for computer-aided detection of tuberculosis," *IEEE Transactions on Medical Imaging*, vol. 35, no. 4, pp. 1013–1024, 2016.

[56] S. Jaeger, A. Karargyris, S. Antani, and T. George, "Detecting tuberculosis in radiographs using combined lung masks," in *Proceedings of the 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4978–4981, San Diego, CA, USA, August 2012.

[57] M. Thiyagarajan and N. Bharathi, "Lung cancer detection using fuzzy auto-seed cluster means morphological segmentation and svm classifier," *Journal of Medical Systems*, vol. 40, p. 06, 2016.

[58] K. S. Mithra and W. R. Sam Emmanuel, "Gfnn: Gaussian-fuzzy-neural network for diagnosis of tuberculosis using sputum smear microscopic images," *Journal of King Saud University - Computer and Information Sciences*, vol. 33, no. 9, pp. 1084–1095, 2021.

[59] K. S. Mithra and W. R. Sam Emmanuel, "Gaussian model based hybrid technique for infection level identification in tb diagnosis," *Journal of King Saud University - Computer and Information Sciences*, vol. 33, no. 8, pp. 988–998, 2021.

[60] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, 1973.

[61] H. Aidos and A. Fred, "k-nearest neighbor classification using dissimilarity increments," *Lecture Notes in Computer Science* in *Proceedings of the International Conf. on Image Analysis and Recognition*, vol. 7324, pp. 27–33, Aveiro, Portugal, June 2012.

[62] P. H. Swain and H. Hauska, "The decision tree classifier: design and potential," *IEEE Transactions on Geoscience Electronics*, vol. 15, no. 3, pp. 142–147, 1977.

[63] S. Xu, "Bayesian Naïve Bayes classifiers to text classification," *Journal of Information Science*, vol. 44, no. 1, pp. 48–59, 2018.

[64] M. Dorfer, R. Kelz, and G. Widmer, "Deep Linear Discriminant Analysis," 2016, https://arxiv.org/abs/1511.04707.

[65] T. Kam Ho, "Random decision forests,"vol. 1, pp. 278–282, in *Proceedings of the 3rd international conference on document analysis and recognition*, vol. 1, IEEE, Montreal, QC, Canada, August 1995.

Hindawi
Journal of Healthcare Engineering
Volume 2022, Article ID 4055491, 14 pages
https://doi.org/10.1155/2022/4055491

Research Article

A Liver Damage Prediction Using Partial Differential Segmentation with Improved Convolutional Neural Network

B. Sumathy,[1] Pankaj Dadheech,[2] Monika Jain,[3] Ankur Saxena,[4] S. Hemalatha,[5] Wenqi Liu,[6] and Stephen Jeswinde Nuagah[7]

[1]*Department of Instrumentation and Control Engineering, Sri Sairam Engineering College, Chennai, India*
[2]*Department of Computer Science & Engineering, Swami Keshvanand Institute of Technology, Management & Gramothan, Jaipur, Rajasthan, India*
[3]*Department of Electronics & Communication Engineering, ITS Engineering College, Greater Noida, Uttar Pradesh, India*
[4]*Indus Institute of Information & Communication Technology, Indus University, Ahmedabad, Gujarat, India*
[5]*Department of Computer Science and Engineering, Panimalar Institute of Technology, Chennai, Tamil Nadu, India*
[6]*Henan Chuitian Technology Co., LTD, Hebi 458000, China*
[7]*Department of Electrical Engineering, Tamale Technical University, Tamale, Ghana*

Correspondence should be addressed to Stephen Jeswinde Nuagah; jeswinde@tatu.edu.gh

Received 2 November 2021; Revised 20 December 2021; Accepted 17 January 2022; Published 27 February 2022

Academic Editor: Waliullah Khan

Background. The liver is one of the most significant and most essential organs in the human body. It is divided into two granular lobes, one on the right and one on the left, connected by a bile duct. The liver is essential in the removal of waste products from human food consumption, the creation of bile, the regulation of metabolic activities, the cleaning of the blood by sensitizing digestive management, and the storage of vitamins and minerals. To perform the classification of liver illnesses using computed tomography (CT scans), two critical phases must first be completed: liver segmentation and categorization. The most difficult challenge in categorizing liver disease is distinguishing the liver from the other organs near it. *Methodology*. Liver biopsy is a kind of invasive diagnostic procedure, widely regarded as the gold standard for accurately estimating the severity of liver disease. Noninvasive approaches for examining liver illnesses, such as blood serum markers and medical imaging (ultrasound, magnetic resonance MR, and CT) have also been developed. This approach uses the Partial Differential Technique (PDT) to separate the liver from the other organs and Level Set Methodology (LSM) for separating the cancer location from the surrounding tissue based on the projected pictures used as input. With the help of an Improved Convolutional Classifier, the categorization of different phases may be accomplished. *Results*. Several accuracies, sensitivity, and specificity measurements are produced to assess the categorization of LSM using an Improved Convolutional classifier. Approximately, 97.5% of the performance accuracy of the liver categorization is achieved with a 94.5% continuous interval (CI) of [0.6775 1.0000] and an error rate of 2.1%. The suggested method's performance is compared to that of two existing algorithms, and the sensitivity and specificity provide an overall average of 96% and 93%, respectively, with 95% Continuous Interval of [0.7513 1.0000] and [0.7126 1.0000] for sensitivity and specificity, respectively.

1. Introduction

Cancer has risen to become one of the most prevalent causes of death in contemporary times, and liver cancer has risen to become one of the three most lethal illnesses in the world throughout time [1]. Oncologists may use segmented liver tumor to confirm changes in tumor size that have occurred. The data may subsequently be used to gauge the patient's reaction to treatment and, if necessary, to offer medical help to the patient. Many applications of a medical image recovery system rely on the classification of medical pictures, which is one of the most important things to consider. When highly varied medical picture data become available, reliable classification algorithms are essential to make appropriate decisions. The CT paradigm is used in clinical diagnostics, which allows radiologists to accurately detect and follow

changes in the body's physiological state over time. On CT pictures, distinct tissues in separate lines with varying gray rates may be distinguished, and this information can be used to make a medical diagnosis.

Cancer is the leading cause of death in numerous industrialized countries, particularly in the United States [2]. Cancer diagnosis in general practice is based on scientific and histological information that may be faulty or wrong, leading to inaccurate conclusions. When you look at the human body from the inside out, you will see that the liver is located in the upper abdomen. The liver's purpose is to absorb and eliminate waste from the blood. When there is an excess of waste cells in the liver, a lump of tissue known as a tumor or growth may develop [3]. A tumor may be either benign or malignant. Because benign malignancies are not carcinogenic, they should be removed by surgeons. Likely, benign tumor will not recur after therapy in most cases. A hemangioma is a benign blood vessel mass that has become twisted and crowded. Cancer is a term used to describe tumors that have become malignant. The majority of primary liver tumors arise in the hepatocytes. Hepatocellular carcinoma and malignant hematoma are used to describe this kind of malignancy. In the early identification and treatment of liver cancer, ultrasonography (US), CT, and magnetic resonance imaging (MRI) are the most often used diagnostic imaging techniques. For detecting a wide range of disorders, including colon cancer, CT is the most used and recommended procedure. CT scans allow a surgeon to confirm the presence of a tumor and assess the size, location, and duration of the tumor with pinpoint accuracy [4]. Thyroid cancer radiotherapy, biopsies, and other minimally invasive treatments may be effectively planned and administered after CT scans.

Smoothing is required throughout the file processing process to make extraction and grading more convenient and accurate. As a result, it is essential to have a flawless filtering approach in biomedical image processing. Selecting the most suitable segmentation algorithm for a liver tumor picture is critical for achieving good performance. The unknown component of the liver pictures may be extracted using a segmentation method that is suitable for the situation. The first two steps in the feature extraction process are preprocessing and segmentation, which incorporate feature collecting as a second phase. When a candidate is chosen, the categorization procedure is carried out. Selecting the most appropriate filter for denoising, segmentation, function selection, and a prediction algorithm for categorizing liver tumor pictures is still a significant research challenge to be completed shortly.

Liver tumors, often known as liver cancers or liver growths, are malignancies or growths that develop on or inside the liver [5]. Hêpar is the Greek word for liver, and many different types of tumor may be found in the liver. Depending on the stage of development, this development might be benign (cancerous) or malignant. If the tumour cell is natural, it will be nice to be around. There was something wrong, and as a result, the situation got disorganised, and a glob was formed. This did not turn out well. They turn into cancer cells when their growth and division become irregular and out of control, and the tumor progresses to cancer.

1.1. Liver Screening. If a person has cirrhosis or another risk factor, it is essentially crucial to follow medical advice, regardless of whether the individual receives frequent liver cancer screenings or not. Early detection will present a higher possibility of success in treating cancer, that is, if discovered before any symptoms appear. Hepatologists are the doctors who have the most significant expertise in screening primary liver cancer. They are also the most expensive. A biopsy, imaging tests such as ultrasound, CT, or CAT scan, or testing for the chemical Alpha-Fetoprotein (AFP) in the blood, which may be produced by cancer cells, are all options for cancer screening (MRI) [6]. More information about these tests may be found in the diagnostic section of the therapy section of this website. According to the etiology of liver illness, various rules must be followed. Diagnosis is the process of identifying an illness based on its symptoms, indicators, and the after-effects of various analytic procedures. The decision reached as a result of the operation as to whether a tumor is malignant or benign is referred to as a diagnostic outcome. When it comes to diagnosing liver cancer, several tests must be performed, and the doctors may do some preliminary tests to see if the disease has moved to another section of the body from where it originated. This condition is referred to as metastasis [7]. A biopsy is the sole technique for a doctor to determine whether or not a specific region of the body is affected by cancer in most cancer cases. A biopsy is a procedure in which a physician removes a tiny sample of tissue to be tested in a laboratory. A biopsy may not be feasible, in which case the doctor may offer additional tests that may aid in the diagnosing process.

The most prevalent primary liver cancer is hepatocellular carcinoma. Hepatocellular carcinoma (HCC) is more common in persons with persistent liver disorders such as cirrhosis from hepatitis B, or C. HCC is often identified without the need for a biopsy [8].

Physicians consider the following variables when deciding on a diagnostic procedure: The kind of cancer suspected, the signs and symptoms experienced by the patient, the patient's age and medical condition, and the findings of previous medical tests.

HCC may be diagnosed with the use of the tests listed below. It is possible to diagnose each form of cancer with the tests on this list; however, not all tests on this list would be recommended for every individual. Examination of the physical body: If a patient exhibits signs of HCC, the physician will palpate the abdomen to look for lumps, swelling, or other abnormalities in the liver, spleen, or other surrounding organs, among other things. The doctor would also examine for indications of jaundice, which include yellowing of the skin and whites of the eyes and an abnormal build-up of fluid in the abdomen. Tests of the blood: In addition to a physical examination, the physician would most likely do a blood test to check for the presence of a chemical known as AFP [9]. AFP is discovered in high

concentrations in the blood of around 50% to 70% of persons with advanced HCC. In addition, the physician would examine the patient's blood to determine whether or not he or she had hepatitis B, or C. Other blood tests may be used to determine how effectively the liver is functioning.

1.2. Medical Imaging Techniques. The development of visual representations of regions in the human body is the medical imaging technology used to diagnose medical issues and monitor treatment outcomes. Zia et al. [10] are the most well-known names in the business. The human body is a highly complex system that requires a great deal of attention. In-depth examination of data's static and changing features adds to the accumulation of large amounts of information. One of the most challenging problems for academics and physicians is figuring out how to gather, store, and display massive amounts of information about the body in a way that can be digested, processed, and utilized to develop more helpful diagnostic tools and treatment processes.

In many cases, presenting information in visuals is the most efficient method for solving this difficulty. We, as human beings, are well aware of this efficiency; we have relied on vision more than any other perceived ability to interact with the world around us from our earliest years. Photomicrographs of a dynamic object, such as the human body, capture elements of the entity, such as its transmission, clarity, emissivity, reflectivity, conductivity, and magnetic properties, together with fluctuations in time in each of these attributes. To offer specifics on the fundamental features of the material, images that depict one or more of these characteristics may be analyzed to provide more information. Physicians are increasingly reliant on such pictures to understand the human body better and intervene in the processes of sickness and damage in the patient. Imaging to organize and analyze biological and medical processes is expected to continue to grow in importance, not just in clinical medicine but also in the biomedical research sector that underpins clinical care and its support. A medical imaging method and the process is a technique and process that is used to take pictures of the human body (or parts and functions of the human body) for clinical reasons (medical treatments that are intended to disclose, diagnose, or analyze illness) or medical science research (including the study of normal anatomy and physiology). It is possible to do medical imaging on excised organs and tissues, although this is not often referred to as medical imaging and is instead considered part of pathology rather than medicine. Computer-Aided Diagnosis (CAD) systems categorize medical images as a fundamental approach. In medical imaging, traditional approaches focus primarily on form, color, texture elements, and their combinations. The majority of these features are issue-specific and complementary in previous studies. As a result, the system is unable to construct representations of high-level issue domain ideas and has a weak model generalization capacity, as seen in Figure 1.

Recent deep learning algorithms give an efficient solution to design an end-to-end model that can calculate final classification labels from the raw pixels of medical pictures

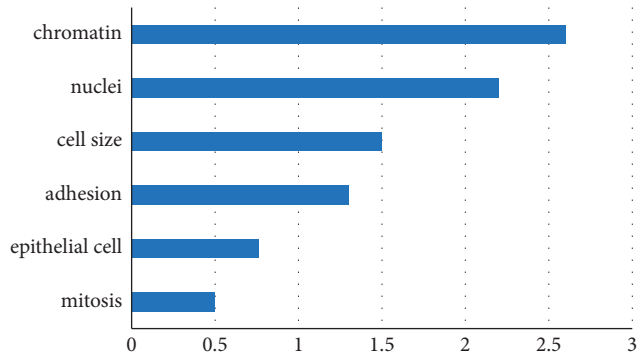


Figure 1: Feature selection concerning liver image.

using just the raw pixels of medical images. On the other hand, deep learning models suffer from high computing costs and constraints in the model layers and channels as a result of the high resolution of the medical pictures and the short dataset size, respectively. A large number of algorithms are being used to tackle these challenges.

Classification is a supervised learning strategy used in deep learning and statistics, in which the computer program learns from the input data and then utilizes this learning to categorize fresh observations. There are many applications for classification issues, including voice and handwriting recognition, biometric identification, document categorization, and illness classification, to name a few [11]. More research is required to increase the accuracy, efficiency, and robustness of liver CT segmentation. The use of enhanced edge detection techniques, parallelization, and the combination of diverse approaches may be the most effective means of developing improvements in liver CT segmentation methods. To correctly identify and segment tumor sections from liver CT scans, even when the pictures are noisy, it is necessary to conduct this study and create unique techniques. During the first phase of this project, the enhanced design was implemented. The segmentation of the liver CT image is accomplished using the ANN approach. This approach is more often used for grouping large amounts of data quickly and effectively. This means that image processing methods, particularly segmentation approaches, may benefit from it. Although this approach does not need any training data, it does require initial segmentation settings, which are the primary benefit of this technique.

A new enhanced residual GoogleNet CNN approach [12] is suggested to increase the performance of the classification process. According to the current study's findings, uniformity in lesion segmentation leads to better results in patients. It has been observed that the present techniques make use of the form specification, which lacks certain qualities such as homogeneity. There are 24 approaches for identifying abnormalities in liver CT scans that are comparable to one another but we only take a restricted range of parameters into account. Such strategies do not improve the performance and accuracy of verification operations.

1.3. Limitations of Existing Approaches. Compared to other techniques, these do not consider several different

categorization functions. The decreased accuracy of the diagnosis and the increased temporal variability made it challenging to detect liver cancers [13]. It is evident from the more rigorously constructed and verified research that machine learning approaches may be utilized to significantly (15–25 percent) enhance the accuracy of predicting cancer susceptibility, recurrence, and death in patients with cancer.

1.4. Contribution of the Proposed Work. Following are the aims of the proposed work:

(i) Develop a unique machine learning soft optimization and deep learning method for liver image classification and segmentation from CT pictures;

(ii) Detect and categorize abnormality in liver CT images; and

(iii) To develop an effective liver tumor detection and classification strategy that considers a more significant number of characteristics.

The research team measured classification accuracy and time complexity to evaluate the liver tumor growth methodologies. Among the achievements made by the current research team is the invention of a filter performance measure for liver image preprocessing.

Biomedical Applications of Intensive Artificial Neural Network (ANN) for Liver Tumor Classification in Health Care [14]

The details of the bulk of the tumor have been improved using the Clinically Visual Deep Neural Network (DNN) technique based on volumetric analysis for liver tumor categorization. The remaining sections of the paper are grouped as follows: The background analysis from the current technique is included in Section 2 of the report. Section 3 contains the technique for segmenting the liver from other organs and the segmentation of liver cancer from the separated liver pictures. Section 4 contains the methodology for segmenting the liver from other organs. Section 3 describes the categorization of the proposed Convolutional Layer, followed by Section 4, which contains the experimental findings and discussion of the proposed Convolutional Layer. Section 5 concludes with a discussion of future work and provides a conclusion.

2. Theoretical Analysis of the Existing Research Methodology

Generally speaking, cancer may be characterized as a condition in which aberrant forms of cells proliferate and spread uncontrollably throughout the body. In medical terminology, the word "neoplasm" (which refers to new growth) should be used rather than the phrase "cancer." "Benign" neoplasms are those that have just the feature of confined growth and are thus categorized as such. A study by Miller et al. was published in 2019 [15]. *Malignant tumors* are defined as those that have the characteristics of invasiveness and the ability to metastasize in addition to their primary characteristics. The word cancer is typically reserved for certain forms of tumor development. Although the term "tumor" really refers to a "local swelling," it is often used interchangeably with the term "cancer" and will continue to be used in this manner in the future.

A benign tumor is often surrounded by fibrous tissue, making surgical excision of the whole tumor relatively straightforward. When examined under a microscope, the cells of a benign tumor seem to be very similar to the cells of the surrounding healthy tissue. Compared to normal cells, which often include a high proportion of cells in a growth-arrested or quiescent stage of the cell cycle, benign tumor cells typically contain an increased proportion of cells in the mitotic (dividing) stage of the cell cycle. The opposite is true for advanced malignant tumors, which, although may be encapsulated in the early stages of their growth, do not have well-defined borders and instead have spread into the surrounding healthy tissues. Furthermore, they may be distinguished from normal and benign tumor cells by their appearance. Not only do they spend most of their time in the mitotic stage of the cell cycle, but they also have an abnormally large number of chromosomes (aneuploidy).

Using a robotized framework for the segmentation and classification of liver tumor, Vadali et al. [16] developed an efficient and straightforward approach. Specifically, the suggested framework includes preprocessing, segmentation, post-processing, and the final categorization into a benign and malignant tumor, among other things. The picture is downsized to a resolution of 256×256 pixels during the preprocessing step. During the segmentation step, the level set approach is linked to segmenting the suspected region. The area of fascination is identified from the initial photograph when it comes to the post-processing step. Finally, the Pseudo Zenerike minute and the GLDM are used to highlight extraction from a CT picture. These components are provided to contribute to the SVM to determine whether the tumor is benign or malignant. The SVM is prepared to make use of four different photos. The suggested framework has an accuracy rate of 86.7 percent, which is rather impressive. A new framework, the CFCSA, was suggested by Anter and Ali [17], in which the crow search algorithm uses the global optimization strategy to overcome the sensitivity of the local optimization technique.

In this approach, the fuzzy c-means (FCM) objective function is employed as a cost function, and the chaotic crow search optimization algorithm is used to find the optimum solution. Benchmarking is performed against the binary crow search algorithm (BCSA), chaotic ant lion optimization algorithm (CALO), binary ant lion optimization algorithm (BALO), and bat algorithm relevant methodologies to see how well the new algorithm CFCSA performs. In this study, the proposed CFCSA algorithm is compared to other algorithms such as the BCSA, CALO, BALO, and bat algorithms. The algorithms are tested in the following areas: diabetes, heart disease, Radiopaedia CT liver imaging, breast cancer, lung cancer, cardiotocography, ILPD, liver disorders, hepatitis, and arrhythmia. According to Diana et al. (2020) [18], the optimal DNN parameters yield the best optimum performance over the datasets under consideration. Wang et al. [19] proposed an efficient sampling strategy based on

Inverse Random Under Sampling (IRUS) to solve the difficulties of class imbalance to improve efficiency.

IRUS undersamples the majority class, resulting in several unique partitions, each of which contains samples from the minority and majority classes separated by a border. An optimization strategy based on the Artificial Plant Optimization (APO) algorithm is also presented to select the most effective and efficient features and parameters of classifiers to increase the effectiveness and efficiency of classification. Using an optimization method, the number of iterations and computation time required for feature selection and parameter selection for classifiers that distinguish between HCC recurrence and nonrecurrence are reduced. Support Vector Machine (SVM) and Random Forest (RF) classifiers categorize patients as having or not having HCC based on optimum characteristics and parameters derived from the data. A review of the principles of deep Improved Convolutional Neural Networks for image classification was published by Qin et al. [20], followed by deep learning to categorize localized hepatic lesions on multiphase CT images of the liver. Nanda et al. [21] investigated deep learning approaches as a first and primary method for extracting the liver from an abdominal CT scan and then, as a result, for segmenting the lesions from a tumor-ridden liver after the liver had been extracted.

To segment lesions, once a tumor has been discovered in the liver by GA-ANN, which has been fed textural liver data using LTEM for its classification method, a cascaded model of Improved Convolutional Neural Networks is utilized. To tackle the segmentation of liver tumor in CT abdominal images, Budak et al. [22] first defined the issue as a classification problem and then solved it using a cascaded classifier architecture based on Deep Improved Convolutional Neural Networks. It was built and taught to detect liver regions and lesions in CT scans with low picture quality using two deep encoder-decoder Improved Convolutional Neural Networks (CNN). In another way, an EDC segments the liver picture and uses the segmentation as input for training a second CNN. Once the tumor areas inside the liver ROI regions have been segmented as anticipated by the first EDC, the second ED CNN may be applied. The segmentation of the hepatic tumor inside the liver ROI also considerably reduces false-positive results.

To quantify the performance of the proposed model, it was tested against a publicly available dataset (3DIRCADb), and many metrics were utilized to assess its performance. Ben-Cohen and colleagues (2018) [23] intended to aid in the identification of liver metastases for CT scans, and it makes use of a completely innovative network (FCN) for both the overall background and local patch detection using a super-pixel sparse classification. The use of CT scans is emphasized unusually. The importance of finding liver metastases, especially tiny metastases, in the early identification of liver cancer cannot be overstated when it comes to the early detection of liver cancer. Ben-Cohen and colleagues (2016) have examined CT exams for liver segmentation and identification of liver metastases, with the entire revolutionary network (FCN) being the most promising. FCN has shown to be quite effective for semi-segmentation. Hawkins et al. have developed a novel machine learning approach to diagnosing HCC in 165 patients, which is based on machine learning (2019) [24].

Maaitah et al. [25] developed an intelligent model for liver illnesses that was built on the Fuzzy Neural System (FNS) and other techniques (FNS). Fuzzy systems and neural networks (FNS) are being investigated to identify liver problems in this context. The FNS's structure and learning method are both given in detail. Wang et al. [19] suggested pretrained deep CNNs on picture patches focused on medical abnormalities and then merged them with class activation mappings and region proposal networks to develop abnormality detectors for medical abnormalities. They were using deep transfer learning algorithms. The findings are compared with the results of three other standard classifier algorithms to increase the classification accuracy of the method. A Deep Improved Convolutional Neural Network was used to develop a technique for identifying liver tumor candidates from CT scans, which was published in 2017. For liver disease screening, Yao et al. [26] suggested a densely connected deep neural network (Dense DNN), which was trained using the most widely used liver function tests (LFTs) and demographic information of individuals.

The suggested approach includes a mechanism for identifying liver tumor while limiting bleeding, particularly on CT images, and when there are a large number of lesions, which is particularly important [27]. According to the current study's findings, uniformity in lesion segmentation leads to better results in patients. It has been observed that the present techniques make use of the form specification, which lacks certain qualities such as homogeneity. It is possible to develop comparable approaches for identifying abnormalities in liver CT scans that merely consider a limited number of features. Such strategies do not improve the performance and accuracy of verification operations. Compared to other techniques, these do not consider several different categorization functions [28]. The decreased accuracy of the diagnosis and the increased temporal variability made it challenging to detect liver cancers.

3. Methodology of Proposed Screening Technique with Various Stages

As a result of the aggressive nature of liver cancer's growth, few treatment choices are available to patients. Many low- and middle-income countries are under tremendous financial strain to provide adequate treatment for people with liver cancer. Having already reliable estimates of liver cancer incidence will assist in concentrating attention on the need to track liver cancer and enable optimal treatment for people diagnosed with the disease [29]. To accurately and timely predict liver cancer, it is necessary to develop suitable prediction methods for the disease.

Figure 2 represents the flowchart of the proposed work. A CT scan is a noninvasive diagnostic imaging technology that uses a mix of X-rays and computer tomography to create horizontal or axial pictures of the

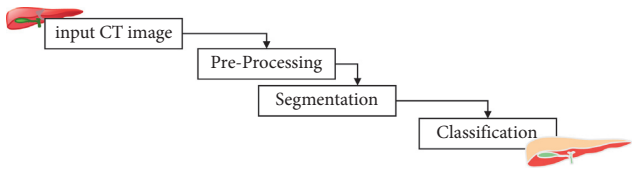


Figure 2: Flow of proposed work.

liver, which may be used to diagnose liver disease [30]. As a result, image processing–based research contributes to the advancement of liver cancer therapy. In computer tomography, the X-ray beam travels around the body in a circle surrounding the patient. It provides a new perspective on the same organ in this way. CT scans may be performed with or without "Contrast," which is a sort of substance that is taken orally and put into an intravenous (IV) line to cause the liver or tissues under observation to become more visible on the screen. The liver is the biggest organ in the body, and it is placed behind the right rib cage and below the base of the lung. It is involved in the digestion of food.

Blood cells are filtered, nutrients are processed and stored, and part of these nutrients is converted into energy [31]. It is also responsible for the breakdown of hazardous chemicals. The left and right lobes of the liver are the two primary hepatic lobes. The quadrate and caudate lobes are visible when the liver is examined below the surface, indicating two extra lobes. Hepatocellular carcinoma (HCC) develops when the liver cells grow out of control and spread to other parts of the body. It is the most common kind of cancer in the United States. Primary hepatic cancers can form when the cells exhibit aberrant activity. According to reports, liver cancer is the second most common disease to cause death in men and the sixth most common cancer to cause death in women. In 2008, around 750,000 individuals were diagnosed with liver cancer, with 696,000 people dying due to the disease. Males get infected at a rate that is twice as high as that of females worldwide. Virus-induced hepatitis may result in the development of liver cancer, which is much more severe. It is estimated that over 1.45 million fatalities occur each year as a result of this virus, according to the World Health Organization [32]. According to the World Health Organization, at 7 percent, Egypt was identified as the nation with the highest prevalence of adult hepatitis C (HIV) infection in 2015. The technique presented in this paper was evaluated using datasets made accessible to the public.

First and foremost, the LiTS dataset was used, which contains 131 CT scan pictures and their corresponding ground facts (clinical annotation). It should be noted that the LiTS dataset also contains a collection of 70 CT scans for testing reasons. However, these pictures do not have any associated annotations [33]. As a result, only the 131 annotated CT scans were considered in this study. There are 24 photos accessible in JPEG format, which were all taken from a DICOM file with dimensions of 630630 pixels and a bit depth of 24 bits.

3.1. Noise Reduction Using Pre-Processing. The preprocessing stage of the diagnosis of liver cancer is the first step in the procedure. Preprocessing is necessary to guarantee the long-term viability and usage of a database. Any step seems to be critical in image processing workflow to achieve this. Filters and histogram equalization methods are used to preprocess unneeded error identification before the actual error detection takes place. Noises may be eliminated using a CT image available at this stage. In nonlinear optical filtering systems, the Adaptive Median Filter (AMF) is often used to remove noise from images or signals, and it is also known as the adaptive median filter. A noise reduction approach is a standard preprocessing method for increasing efficiency. When a CT scan is taken, preprocessing is done to improve the picture's contrast. Histogram equalization is often used to improve picture consistency by balancing the histograms.

Histogram Equalization is a computer-assisted procedure used to improve the contrast of photographs. The most common sensitivity values are significantly improved, i.e., the picture intensity range is significantly widened. It makes it possible to reduce local contrast to improve ties between areas. As a result, after applying the histogram equalization, the average contrast of the photos is enhanced by a significant amount. Equation (1) represents the intensity increment in the input image

Let q denote the normalized histogram of each possible intensity. Hence,

$$q^{\underline{x}} = \left(\text{Intensity of } \frac{\text{image}}{\text{total}} \text{ number of pixels}\right). \tag{1}$$

The histogram equalized image can be defined as

$$Z_{x,y} = \text{base}\left((Z-1)\sum_{y=0}^{c_4} xq^z\right), \tag{2}$$

where the base is the integer that is closest to the given value. This is the same as converting the pixel intensity but in reverse.

$$\frac{\partial O}{\partial y}\left(\int_0^O yqO(y)da\right) = \partial O(O)\left(y^{-1}\right)(O)\frac{d}{dO}. \tag{3}$$

This is where, at long last, the probability distributed uniformity function may be expressed as $\partial O/\partial y$.

Equations (2) and (3) represent the preprocessing of the input image. While the results show that the equalization procedure utilized produces flat histograms, it may also soften and enhance the appearance of histograms.

3.2. Liver Segmentation Using Partial Differential Technique. In recent years, it has become an increasingly critical and time-consuming effort to segment both the liver and the tumor area. This is accomplished by developing multiple segmentation methods that segment the picture using different modalities. The approaches used for segmentation are classified into three categories: manual segmentation, semi-automatic segmentation, and fully automated segmentation. Medical specialists do manual segmentation of images layer by layer. The liver borders are detected by two distinct

radiologists or by the same radiologists at various times in a single picture when manual segmentation is performed. The structure and observation of the medical pictures are explained using this sort of a segmentation method. There are many issues in this area, the most significant of which is reducing picture quality and the increase of artifacts. Equation (4) represents the input segmentation from another organ in the image. L be the input pixel, and O be the intensity of the image.

$$L[y] = O[y] + S[o] + \phi, y \in S, \tag{4}$$

subject to the initial conditions

$$v(y,0) = h_0(y)_+ v_t(y,0) = h_1(y), \ldots, v_t^{(x-1)}(y,0) = h_{-1}(y), \tag{5}$$

and spatial conditions are represented in

$$u(\theta,t) = h_0(t),\ u_z(0,t) = \hbar_1(t), \tag{6}$$

where L is the oth order derivative w.r.t "t," $R[1]$ is a linear operator and NH is the nonlinear operator with degree three. And, $\phi = \phi(x,t)$ and $u = u(x,t)$ are two types of functions: known and unknown. Manual segmentation has many significant problems, the most significant of which are the huge number of picture slices required, the lengthy time required, and the lack of satisfactory results. Furthermore, the process of creating a separate dataset is a time-consuming and difficult one. A semi-automatic segmentation system has been created in response to these challenges, which interactively determines the seed sites for determining the liver's border. As a result of this method, we first apply the Laplace transformation to equation (7) concerning the variable t, and we get

$$v(y) = L[O[v(y)] + S[v(y)]] + L[\phi(y,x)]. \tag{7}$$

By using I.C. (7), we get

$$t^o \tilde{v}(y,t) = i(y,t) + L[O|v(y,x)] + G[v(y,x)] + \tilde{\phi}(y,t), \tag{8}$$

where

$$\overrightarrow{i}(y,t) = \sum_{s=0}^{1} x t^{(s+1)} \mu_m^2, \tag{9}$$

and $u(y,x)$ and $\overrightarrow{\phi}(y,x)$ are the Laplace transformed forms of $u(y,x)$ and $\phi(y,x)$, respectively.

Now, dividing by t^o on both sides, we get

$$u(x,s) = \left[\left[\frac{h(x,s)}{s^n} + \frac{1}{s^n} L[Nu(x,t)] + R[tu(x,t)]\right]\right] + \frac{\tilde{\phi}(y,x)}{t^o}, \tag{10}$$

where x, y represent the pixels rows and columns, respectively, with the intensity level and smoothening range as t

$$\tilde{f}(y,t) = \frac{i(y,t)}{t^o} + \frac{o(x,s)}{s^n}. \tag{11}$$

Now, we apply inverse Laplace transformation on equation (11) concerning 8^7, and then we get

$$Vx = C^{-1}[\tilde{f}(x,y)] + C^{-1}\left[\frac{1}{y^o} L[Ov(y,u)] + S[v(y,t)]\right]. \tag{12}$$

In the second step, we apply a differential transformation on equations (11) and (12) concerning $^4x'$, and we get and

$$V_0(u) = h_0(u),\ U_1(u) = i_1(u), \tag{13}$$

where $V_u(v)$ and $G_l(v)$ are the differential transforms of $v(x,y)$ and $z(x,y)$, respectively The closed-form of the solution can be expressed as follows using the aforementioned recurrence equation and initial conditions.

Using this method, it is simple to forecast the properties of the picture being created. The segmentation approach in this work is based on gray scales, which are employed in conjunction with a computer. Specifically, the difference in size between big and tiny pixels surrounding the object's edges is assessed.

$$X^{\text{segment}} = \sum_{s}^{o} xV_2\left(V_x, V_y\right) \cdot Vo \cdot log_{c_i} + \gamma \int c_i dy, \tag{14}$$

where $W^{\text{Segmentation}}$ is the watershed segmentation, $V2$ is the velocity gradient, V_x, V_y. When a pixel value is low or high, the image's spatial size is represented by log (c_i). When a frequency coefficient is represented by a distance between pixels, the image's spatial size is represented by log (c_i).

3.3. Liver Cancer Cells Segmentation Using Level Set Methodology. The conclusion of the local segmentation had been improved with the ith slice of all n slices, for more precision, and with the use of the true benefits of the Force Function (FF) function, which had been implemented. Geodesic Active Contour (GAC), as well as the Chan-Vese (CV) models, is used in this modification. There are values for the SPF function that are in the range [−1,1]. It altered the signs of the pressure forces within and outside the zone of intersection, causing the contour to shrink when the item was outside the intersection region and to grow when the object was inside the intersection region.

In equation (15), the SPF function is created in the following manner

$$fn(j(y)) = \frac{v(y) - d_1 + d_2/2}{n((y) - d_1 + d_2/2)}, y \in \Omega, \tag{15}$$

open function of H^2, $j(x)$ is the given image in Ω, d_1 and d_2 are defined in equations (16) and (17), respectively.

$$d_1(\phi) = \frac{\int_\Omega (y) I(\phi) dy}{\int_i I(\phi) dy}. \tag{16}$$

Here, in the Heaviside function, $I(\phi)$ is approximated by a smoothed functional I_F which is defined by equation (17).

$$I_d(A) = \frac{1}{2}\left(1 + \frac{2}{\pi}\arctan\left(\frac{\pi}{j}\right)\right). \quad (17)$$

The significance of equation (17) may be described in the following manner. Even if the intensities within and outside the object are homogeneous and c 1c 2, it is evident that Min (Jy)d1, d2 Max (J y), and the equal signs cannot be reached simultaneously everywhere contour, which dominates, is present. As a result, we get the following equation:

$$\text{Min}\left(1(x) < \frac{d_1 + d_2}{2}\text{Max}1(y)\right), y \in \Omega. \quad (18)$$

When the object is in the opposite position, the function is defined by substituting the fn function, in Equation 18 in the level set formulation, we derive as in the following:

$$\frac{\partial\phi}{\partial t} = fn(j(y)) \cdot \left(\text{div}\left(\frac{\nabla\phi}{[\nabla\phi)} + \alpha\right)|\nabla\phi 1 + \nabla fnm(y)\right) \cdot \nabla\phi, y \in \Omega. \quad (19)$$

A gaussian filtering technique is used to further regularize the degree of regularity to prevent the re-initialization procedure. div(/(||))|| is used to represent Gaussian vectors, which makes the term $fn(I(x))$ in equation (19) superfluous. As a result, the level set formulation may be simplified in equation (20)

$$\frac{\partial\phi}{\partial t} = spf(I(x)) \cdot a|\nabla\phi|, x \in \Omega, \quad (20)$$

The finely segmented liver pictures, referred to as slices, are the result of a level set process that ensures correct closeness on both sides of the liver image.

3.4. Feature Extraction and Classification of Liver Cancer Cells. Following the segmentation step, the Gray Level Co-occurrence Matrix (GLCM) may be used to pick the characteristics that are of interest. When used in conjunction with the Gray Level Co-occurrence Matrix approach (GLCM), it is possible to derive second-order statistical texture attributes. The approach has been employed in a variety of applications, and the presence of three or more pixels may be noticed in the third and higher-order textures, indicating that the technique is being used.

Figure 3 represents the feature extraction of DVW. Because it is an arithmetical function, the GLCM can effectively eliminate artifacts in most cases. It is also possible to maintain the accuracy of the picture. It is possible to extract the picture for use in the research process. GLCM can determine the frequency of the pixels with a defined level of accuracy. The single-pixel in issue is to be questioned here, and another pixel is to be referred to as the l route and the neighboring value detachment of m, which are two different pixels. Ordinarily, m only acquires a single value and may profit either way. The generated directional value may then be used to eliminate the properties of the pictures that were utilized in the segmentation procedure.

$$Q(n,t) = H(n,t,p,\phi)\sum_{n=1}^{i} t \sum_{t=1}^{i} tH(n,t,p,\phi). \quad (21)$$

G represents the frequency vector, m represents the frequency of the particular component that will generally have the pixel values of n, and H, p represent the normalized constant, where P represents the features of an image, (n, t) represents the component of the n, and H, p represent the normalized constant.

Improved Convolutional Neural Network (ICNN) is a neural network that performs the classification and grouping of pictures quickly and efficiently. The raw pixels of cell photos are used to build the highlight representations in this approach, and the pixel is acquired by using a hand-created system as in previous ways. In addition, the characterization layer is mutually advanced with these component descriptions to predict the class for each cell image in the database. The following opinions are included in the present study and broadened: A more detailed and point-by-point representation of the ICNN classifier structure is shown. Various critical components for putting together this structure are discussed and provisionally considered. This paper presents three main findings:

Figure 4 represents the classifier structure. The role of pivoting cell pictures in information growth is dissected from top to bottom. The suitability of cell picture veils for this grouping assignment is examined, and an illustration of the astounding versatility of the ICNN-based characterization framework to various datasets is provided. Furthermore, further exploratory connections between the ICNN-based structure and the cutting-edge hand outlined shallower grouping models are aimed to illustrate the focus points of the ICNN classifier system–based cell picture classification, notwithstanding the above. When it comes to producing tests based on learning references after sparsity has been determined, the ICNN yield, which includes the sparsity level and records of initial coefficients in insufficient vectors, are used as the two arrangements of components for creating the NN. The recommended diminutive size ICNN grouping is applied in this approach, which uses 3D stomach CT scans in a prepared pipeline to natural criminologists. The International Convention on Nuclear Nonproliferation is characterized as follows:

$$NZC = \sum_{i}^{1} x\|y_i - D_{xi}\|2 \le OjRMSE^2. \quad (22)$$

For each dictionary set D, ND denotes the number of sparsity levels present, NZC the intended sparsity level and an average number of nonzero coefficients, and RMSE the reconstruction of error or root-mean-square

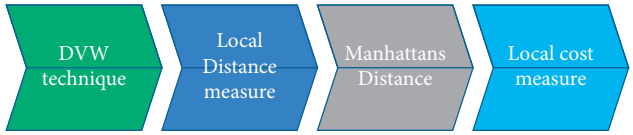


Figure 3: Feature Extraction using Dynamic Vector Warping (DVW).

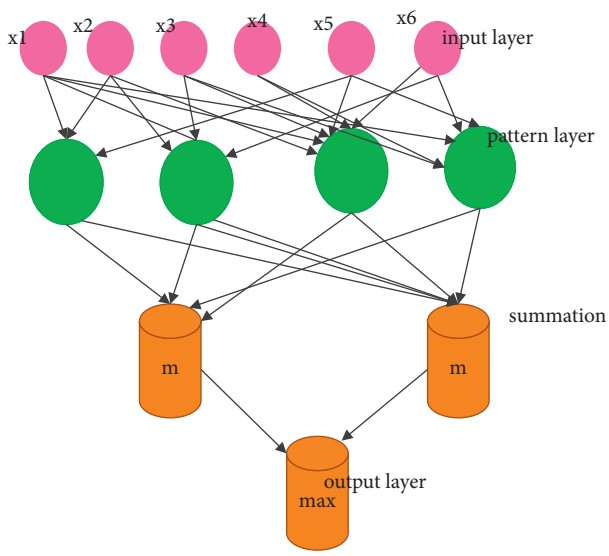


Figure 4: Proposed improved convolutional neural network.

error, respectively. It is input into the ICNN classifier, and the weights are changed until the ICNN returns results that are less than 2.0 for benign instances of liver tumor and more than 2.0 for malignant cases of liver tumor, respectively.

4. Experimental Analysis of Proposed Work

The proposed work is implemented using MATLAB software with training and testing images. In this step, two photos of malignant liver tumor are taken from the datasets that have been utilized.

CT scanning has risen to become an essential imaging technique in diagnosing liver tumor. Different studies have utilized different methodologies for classifying liver development based on CT.

Figure 5 represents the segmentation outcomes of the proposed work. The accuracy, sensitivity, and specificity of the suggested approach have been examined, as well as its overall performance

Figure 6 represent the proposed liver segmentation techniques. According to relevant features, the number of liver cancers that are accepted in the result when compared to other tumors, whereas specificity (irrelevant features) refers to the number of nonliver tumors that are rejected in the outcome when compared to other tumors.

The overall accuracy is applied to the end performance, whereas the sensitivity is applied to the acceptance capacity.

The equation is given below.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{23}$$

$$\text{Specificity} = \frac{TN}{TN + FP},$$
$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}, \tag{24}$$
$$\text{PPV} = \frac{TP}{TP + FP}.$$

(i) True Negative (TN): When it is (F) the samples can be classified or false (F)

(ii) False Positive (FP): When it is (F) the samples can be cl + assified as (T)

(iii) False Negative (FN): when it is (T) the sampled can be classified as (F)

To evaluate how well a classification system performs, accuracy is the most widely utilized parameter. The degree to which classifying rules square measure accuracy determines a classification system's accuracy. Table 1 shows the feature value of the images.

The values are shown in parentheses. The categorization is based on the values of these feature attributes. Following that, the outputs of the feature extraction vectors are trained, and the accuracy of the performance using different classifier approaches is compared. As shown in Table 2, the sensitivity, specificity, and accuracy of the different classifiers are tested in terms of the

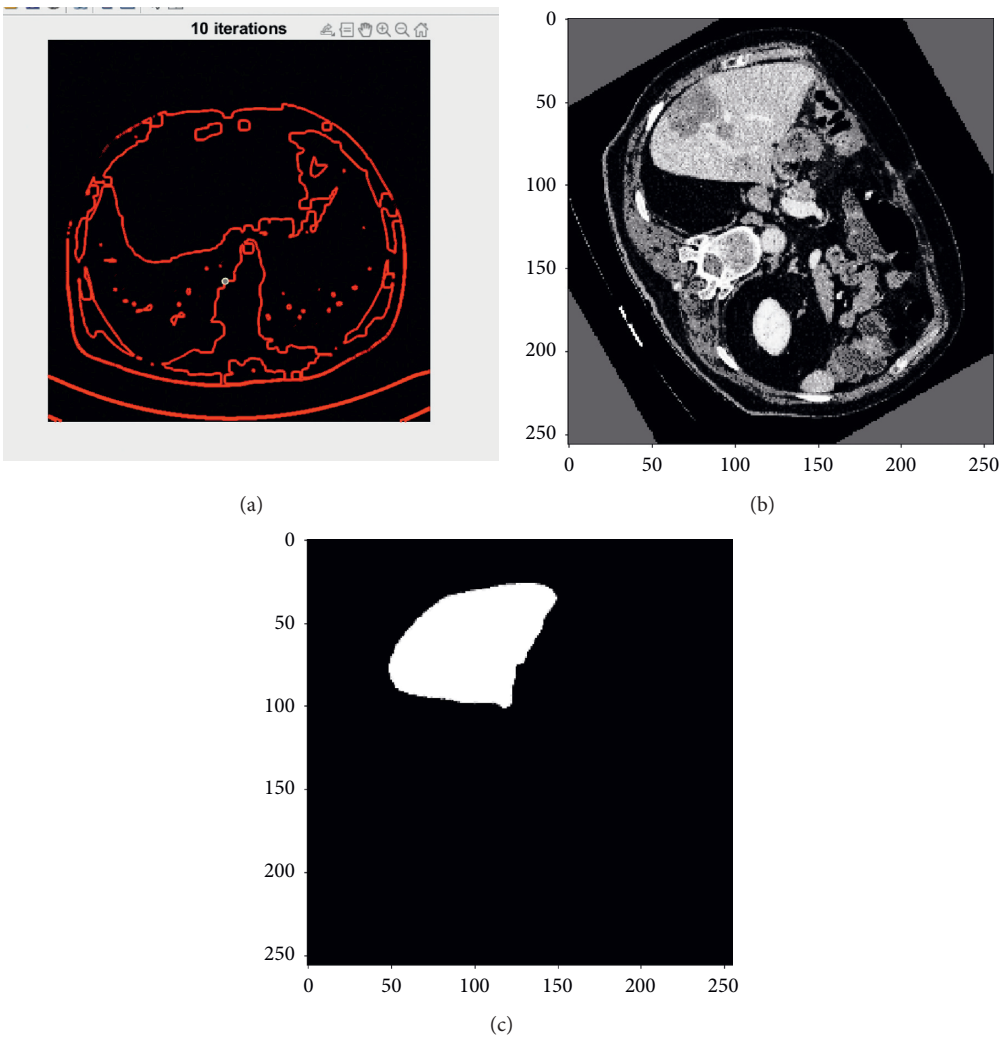


Figure 5: Partial Differential Technique with Input CT image of liver cancer cells.

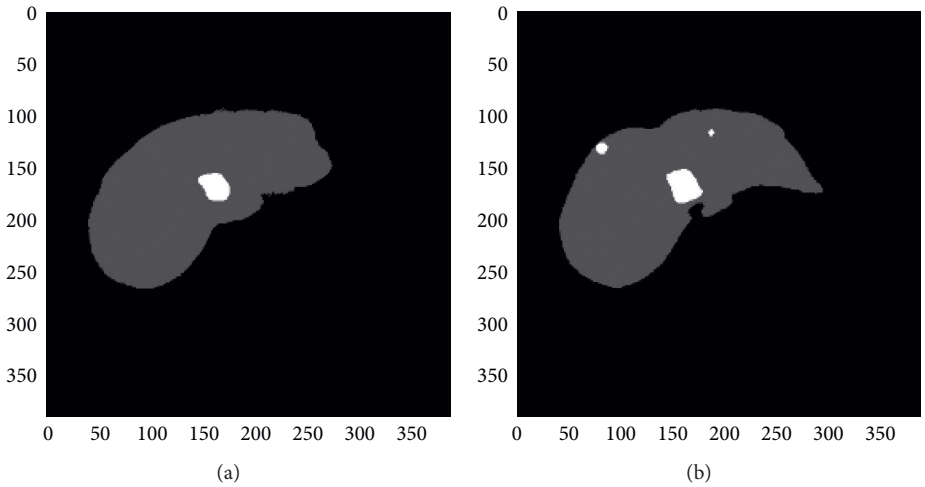


Figure 6: Level Set Technique with Input CT image of liver cancer cells.

Table 1: The feature values that were picked for 7 images.

Image	DVW	Local distance measure	Local cost measure	Entropy
1	858.000000	858.000000	858.000000	858.000000
2	84.000000	26.820513	0.040793	0.086247
3	32.000000	8.497948	0.197925	0.280892
4	25.000000	13.000000	0.000000	0.000000
5	13.000000	20.000000	0.000000	0.000000
6	8.497948	25.000000	0.000000	0.000000
7	26.820513	32.000000	0.000000	0.000000

Table 2: The classification results and performance with the existing classifiers.

Techniques	Accuracy	Sensitivity	Specificity	ROC
SVM	92.36	94.6	92.6	96.6
KNN	93.89	95.66	93.66	97.66
Naïve Bayes	94.36	95.98	96.98	98.98
CNN	95.6	95.36	97.36	97.36
ICNN	97.5	96	93	95

sensitivity, specificity, and accuracy using a variety of classifiers.

Approximately, 97.5% of the performance accuracy of the liver categorization is achieved with a 94.5% Continuous Interval (CI) of [0.6775 1.0000] and an error rate of 2.1%. The suggested method's performance is compared to that of two existing algorithms, and the sensitivity and specificity provide an overall average of 96% and 93%, respectively, with 95% Continuous Interval of [0.7513 1.0000] and [0.7126 1.0000] for the sensitivity and specificity

Figure 7 represents the performance metrics. When comparing the suggested technique to the current methods, the new method achieves superior results in terms of sensitivity, specificity, and accuracy. The accuracy of the classifier is used to estimate how successful it is by displaying the proportion of correct answers. Classifiers such as the ICNN and IANN perform much better in this situation than other classifiers. As a result, the ICNN classifier is capable of successfully classifying a greater amount of data than the other classifiers. The sensitivity and specificity of the classifier are used to evaluate its efficacy.

Figure 8 depicts the classification accuracy comparison results, with the proposed ICNN approach outperforming the other methods in terms of classification accuracy. The term "sensitivity" refers to how well a classification system performs regarding the number of adequately diagnosed benign tumors. In contrast, "specificity" refers to how well the classification system performs regarding the number of correctly classified malignant tumors. Compared to the other classifiers, the ICNN obtains a better level of sensitivity. Given that the goal of cancer detection is to determine whether a patient has cancer or not, which is indicated by the presence or absence of malignant tumors, the most fantastic accuracy and specificity are more significant in research. The

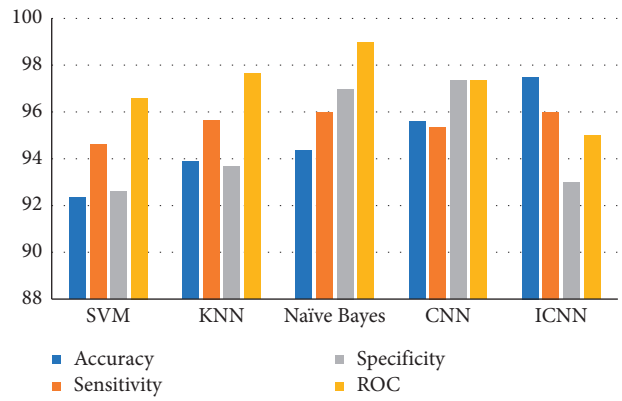


Figure 7: Performance metrics.

rationale for this is that individuals diagnosed with cancer may be further studied to extend their lives, but those categorized as usual would stay unnoticed.

Figure 9 includes the ROC curves that are commonly used to evaluate the performance of binary classification algorithms, and they are also known as receiver operating characteristic curves. The performance of a classifier is represented graphically, rather than as a single numerical number, as is the case with most other metrics.

The computer vision field is built on the concepts of picture recognition and image generation. Despite the fact that both are developing domains, specific approaches from both subareas may sometimes create a dichotomy. Historically, the topic of deep learning (DL) was extensively popularized in discriminative image classification with the AlexNet architecture and picture synthesis with GANs and Variational Autoencoders, among other applications [34].

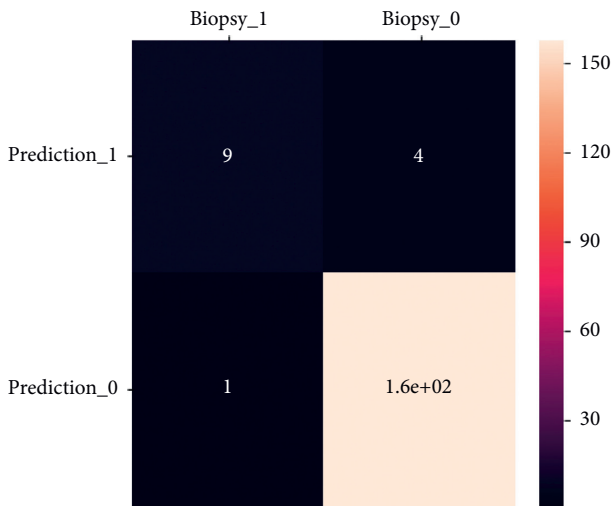


Figure 8: Confusion matrix obtained from training stage of input classifier.

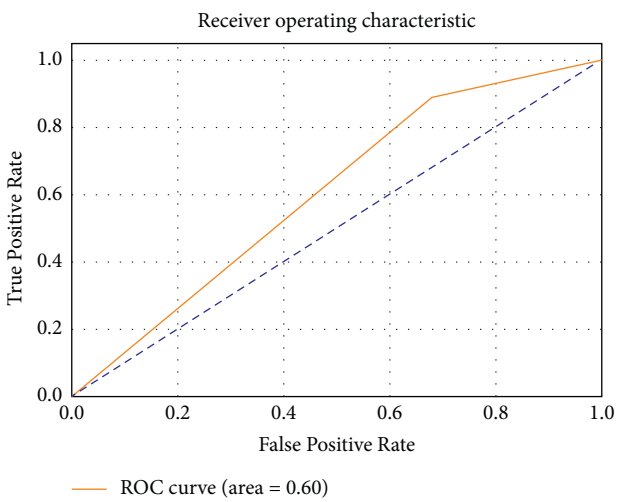


Figure 9: ROC of input classifier.

5. Conclusion and Future Work

Using abdominal CT medical images, a region-based image segmentation approach is provided to identify and segment the liver, which is the area of interest, as the region of interest. The suggested segmentation approach uses a local segmentation step, which provides an approximate boundary for the area of interest in the segmented region of interest. The level set approach fine-tunes this approximate border of the liver even further by using a region-based force function, which effectively stops the contours at weak or hazy margins. The algorithm's efficiency is measured using a variety of metrics and then compared to other conventional algorithms to see which is more efficient. Other soft computing approaches, such as image classification algorithms, may be investigated in the future to lower the computation processing time of liver tumor pictures. Testing against a variety of database pictures and multiple classifier models and optimization methodologies may help improve the accuracy of the liver tumor classification system. Other categorization parameters for liver tumor images are also being evaluated for inclusion in the classification process. Images of the same patient may be evaluated using a variety of imaging modalities, including ultrasound and 0.

To increase the accuracy of the classifier, hybrid classifier approaches may be used. It may be necessary for the liver tumor classification method to evaluate a more significant number of picture samples.

Appendix

A. Partial Differential Technique

When the number of labels is zero, the following is true:

S: labels 1 to N are used for markers, while labels $N+1$ to NO are used for seed.

Place a marker for each area labelled from 1 to N on the map.

All pixels should be sorted in ascending order.

For each pixel, begin by determining the number of labels that exist in N. (z)

If the number of labelled pixels is zero, the current pixel obtains a new label, resulting in the creation of a new temporary area.

Alternatively, Alternatively, if (labelled pixel = 1), the current pixel is assigned this label.

Alternatively, Alternatively, if (labelled pixels ≥ 2), the current pixel is assigned an edge label.

Alternatively, Alternatively, if (labelled pixels is a negative number)

In order to merge the broken pictures and leaky images, the linear interpolation concept is used to interpolate them in order to locate the closest neighbor pixels. (It is also possible to wear a Mexican hat.)

Else \s

begin

Combine all of the neighbors into a single area; the current pixel receives this name.

End \s

End

B. Level Set Methodology. Initialize:

Find the liver tumor cells interior blocks

While (elements of ϕ is interchanged in the pixel values)

For $i=1$ to input number of speed iterations

For each pixel x in $I_d(A)$

If $(d_1(\phi)==1)$ switch x from $I_d(A)$ to $I_d(A')$

Clean $I_d(A')at\ x$

For $i=1$ to output curvature selection

For each pixel x in $I_d(A)$

If $((\phi(x))==1)$ switch x from $I_d(A\prime)$ to $I_d(A')$

Clean $I_d(A')$ at x
End
end

Data Availability

The data that support the findings of this study are available on request from the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] T. Yamashita and X. W. Wang, "Cancer stem cells in the development of liver cancer," *Journal of Clinical Investigation*, vol. 123, no. 5, pp. 1911–1918, 2013.

[2] A. Jemal, F. Bray, M. M. Center, J. Ferlay, E. Ward, and D. Forman, "Global cancer statistics," *CA: A Cancer Journal for Clinicians*, vol. 61, no. 2, pp. 69–90, 2011.

[3] H.-Y. Chen and E. White, "Role of autophagy in cancer prevention," *Cancer Prevention Research*, vol. 4, no. 7, pp. 973–983, 2011.

[4] M. Chand, D. S. Keller, H. M. Joshi, L. Devoto, M. Rodriguez-Justo, and R. Cohen, "Feasibility of fluorescence lymph node imaging in colon cancer: flicc," *Techniques in Coloproctology*, vol. 22, no. 4, pp. 271–277, 2018.

[5] X. Chen, S. T. Cheung, S. So et al., "Gene expression patterns in human liver cancers," *Molecular Biology of the Cell*, vol. 13, no. 6, pp. 1929–1939, 2002.

[6] M. H. Attwa and S. A. El-Etreby, "Guide for diagnosis and treatment of hepatocellular carcinoma," *World Journal of Hepatology*, vol. 7, no. 12, p. 1632, 2015.

[7] J. Bruix and M. Sherman, "Management of hepatocellular carcinoma," *Hepatology*, vol. 42, no. 5, pp. 1208–1236, 2005.

[8] D. C. Rockey, S. H. Caldwell, Z. D. Goodman, R. C. Nelson, and A. D. Smith, "Liver biopsy," *Hepatology*, vol. 49, no. 3, pp. 1017–1044, 2009.

[9] A. Ko and M. Dollinger, *Everyone's Guide to Cancer Therapy: How Cancer Is Diagnosed, Treated, and Managed Day to Day*, Andrews McMeel Publishing, Missouri, United States, 2008.

[10] M. K. Zia, B. Griffel, and J. L. Semmlow, "Robust detection of background noise in phonocardiograms," in *Proceedings of the 2011 1st Middle East Conference on Biomedical Engineering*, pp. 130–133, IEEE, Harjah, United Arab Emirates, February 2011.

[11] M. Faundez-Zanuy, J. Fierrez, M. A. Ferrer, M. Diaz, R. Tolosana, and R. Plamondon, "Handwriting biometrics: applications and future trends in e-security and e-health," *Cognitive Computation*, vol. 12, no. 5, pp. 940–953, 2020.

[12] M. Gour, S. Jain, and T. Sunil Kumar, "Residual learning based CNN for breast cancer histopathological image classification," *International Journal of Imaging Systems and Technology*, vol. 30, no. 3, pp. 621–635, 2020.

[13] S. A. Khan, S. Tavolari, and G. Brandi, "Cholangiocarcinoma: epidemiology and risk factors," *Liver International*, vol. 39, no. S1, pp. 19–31, 2019.

[14] A. D. Wilson and M. Baietto, "Advances in electronic-nose technologies developed for biomedical applications," *Sensors*, vol. 11, no. 1, pp. 1105–1176, 2011.

[15] K. D. Miller, L. Nogueira, A. B. Mariotto et al., "Cancer treatment and survivorship statistics, 2019," *CA: A Cancer Journal for Clinicians*, vol. 69, no. 5, pp. 363–385, 2019.

[16] S. Vadali, G. V. S. R. Deekshitulu, and J. V. R. Murthy, "Analysis of liver cancer using data mining SVM algorithm in MATLAB," *Advances in Intelligent Systems and Computing: Soft Computing for Problem Solving*, Springer, Singapore, pp. 163–175, 2019.

[17] A. M. Anter and M. Ali, "Feature selection strategy based on hybrid crow search optimization algorithm integrated with chaos theory and fuzzy c-means algorithm for medical diagnosis problems," *Soft Computing*, vol. 24, no. 3, pp. 1565–1584, 2020.

[18] A. Diana, F. Carlino, E. Franzese et al., "Early triple negative breast cancer: conventional treatment and emerging therapeutic landscapes," *Cancers*, vol. 12, no. 4, p. 819, 2020.

[19] H. Y. Wang, B. Zhang, J. N. Zhou et al., "Arsenic trioxide inhibits liver cancer stem cells and metastasis by targeting SRF/MCM7 complex," *Cell Death and Disease*, vol. 10, no. 6, p. 453, 2019.

[20] S. Qin, M. C. Wu, Z. Y. Tang et al., "Guidelines on the diagnosis and treatment of primary liver cancer (2011 edition)," *Chinese Clinical Oncology*, vol. 1, no. 1, p. 10, 2012.

[21] N. Nanda, P. Kakkar, and S. Nagpal, "Computer-aided segmentation of liver lesions in CT scans using cascaded convolutional neural networks and genetically optimised classifier," *Arabian Journal for Science and Engineering*, vol. 44, no. 4, pp. 4049–4062, 2019.

[22] Ü. Budak, Y. Guo, E. Tanyildizi, and A. Şengür, "Cascaded deep convolutional encoder-decoder neural networks for efficient liver tumor segmentation," *Medical Hypotheses*, vol. 134, p. 109431, 2020.

[23] A. Ben-Cohen, I. Diamant, E. Klang, M. Amitai, and H. Greenspan, "Fully convolutional network for liver segmentation and lesions detection," in *Deep Learning and Data Labeling for Medical Applications*, pp. 77–85, Springer, Cham, 2016.

[24] M. A. Hawkins, K. K. Brock, C. Eccles, D. Moseley, D. Jaffray, and L. A. Dawson, "Assessment of residual error in liver position using kV cone-beam computed tomography for liver cancer high-precision radiation therapy," *International Journal of Radiation Oncology, Biology, Physics*, vol. 66, no. 2, pp. 610–619, 2006.

[25] M. K. S. Maaitah, R. Abiyev, and I. J. Bush, "Intelligent classification of liver disorder using fuzzy neural systems," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 12, pp. 25–31, 2017.

[26] T. Yau, V. Y. F. Tang, T.-J. Yao, S.-T. Fan, C.-M. Lo, and R. T. P. Poon, "Development of Hong Kong Liver Cancer staging system with treatment stratification for patients with hepatocellular carcinoma," *Gastroenterology*, vol. 146, no. 7, pp. 1691–1700, 2014.

[27] X. Li, J. Li, Y. Liu, Z. Ding, and A. Nallanathan, "Residual transceiver hardware impairments on cooperative NOMA networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 680–695, Jan. 2020.

[28] V. Roy, P. K. Shukla, A. K. Gupta, V. Goel, P. K. Shukla, and S. Shukla, "Taxonomy on EEG artifacts removal methods, issues, and healthcare applications," *Journal of Organizational and End User Computing*, vol. 33, no. 1, pp. 19–46, 2021.

[29] X. Li, M. Zhao, M. Zeng et al., "Hardware impaired ambient backscatter NOMA systems: reliability and security," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2723–2736, Apr. 2021.

[30] P. K. Shukla, J. Kaur Sandhu, A. Ahirwar, D. Ghai, P. Maheshwary, and P. K. Shukla, "Multiobjective genetic algorithm and improved convolutional neural network based COVID-19 identification in chest X-ray images," *Mathematical Problems in Engineering*, vol. 2021, Article ID 7804540, 9 pages, 2021.

[31] X. Li, Y. Zheng, M. D. Alshehri et al., "Cognitive AmBC-noma IoV-MTS networks with IQI: reliability and security analysis," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12, 2021.

[32] S. Pandit, P. K. Shukla, A. Tiwari, P. K. Shukla, and R. Dubey, "Review of video compression techniques based on fractal transform function and swarm intelligence," *International Journal of Modern Physics B*, vol. 34, no. 8, Article ID 2050061, 2020.

[33] G. Khambra and P. Shukla, "Novel machine learning applications on fly ash based concrete: an overview," *Materials Today Proceedings*, pp. 2214–7853, 2021.

[34] A. B. Tufail, Y. K. Ma, M. K. A. Kaabar et al., "Deep learning in cancer diagnosis and prognosis prediction: a minireview on challenges, recent trends, and future directions," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 9025470, 2021.

Hindawi
Journal of Healthcare Engineering
Volume 2022, Article ID 8128689, 7 pages
https://doi.org/10.1155/2022/8128689

Research Article

Early Warning Model of Placenta Accreta Spectrum Disorders Complicated with Cervical Implantation: A Single-Center Retrospective Study

Siming Xin,[1] **Hong Wan**,[1] **Xiaoming Zeng**,[1] **Yanyan Fu**,[2] **Zhizhong Wang**,[1] **Hua Lai**,[1] **Ying Xiong**,[1] **Jiusheng Zheng**,[1] **and Lingzhi Liu**[1]

[1]*Department of Obstetrics, Maternal and Child Health Hospital Affiliated to Nanchang University, Nanchang 330006, Jiangxi, China*
[2]*Department of Preventive Health Care, Maternal and Child Health Hospital Affiliated to Nanchang University, Nanchang 330006, Jiangxi, China*

Correspondence should be addressed to Jiusheng Zheng; zjsheng2012@sina.com and Lingzhi Liu; liulzhi2007@sina.com

Received 10 December 2021; Accepted 13 January 2022; Published 4 February 2022

Academic Editor: Han Wang

Background. Placenta accreta spectrum (PAS) disorders seriously threaten the safety of the mother and infant in the perinatal period. Moreover, PAS is associated with poor maternal and perinatal outcomes once complicated with cervical implantation. Dismally, there are few reports about PAS complicated with cervical involvement currently, and the early warning models are also rarely reported. To screen the risk factors of PAS complicated with cervical implantation and construct an early risk warning model, we performed the analysis of clinical indicators and images of PAS patients by artificial intelligence (AI) data processing methods. *Methods*. The clinical data of 166 patients with PAS in our hospital from January 2016 to September 2020 were retrospectively analyzed. The patients were divided into cervical implantation group and lower uterine implantation group according to the position of placenta implantation. Then, we compared the pregnancy outcomes of the two groups, screened the possible related factors of PAS complicated with cervical implantation by univariate analysis, and established the early warning model by logistic regression analysis. *Results*. The maternal outcome of PAS complicated with cervical implantation was worse than that of the lower uterine implantation group. Through univariate analysis and logistic regression analysis, we found that the cervical width, abundant cervical blood flow, and bladder line interruption were all risk factors of PAS complicated with cervical implantation, and their contribution to the establishment of the regression model was statistically significant. *Conclusion*. PAS complicated with cervical implantation was extremely severe. Early identification of risk factors and establishment of a risk warning model have certain guiding significance for clinical formulation of a reasonable treatment plan.

1. Introduction

Placenta accreta spectrum (PAS), a serious obstetric disease, is characterized by the inability of placenta to be delivered by itself after delivery of fetus, and the tight adhesion between placenta and uterine wall during artificial placenta removal. Moreover, some adverse pregnancy outcomes, including severe obstetric bleeding, diffuse intravascular coagulation, and hysterectomy, may be induced once the villous tissue embedded in the myometrium is forcibly stripped [1, 2]. The incidence of PAS in most middle- and high-income countries has almost increased by 10 times, and that in the United States has increased by 10–15 times compared with 30–40 years ago in the context of increased Cesarean section rate and number of elderly pregnant women [1]. China is known to be among the highest Cesarean section rate all over the world. With the full liberalization of the "two-child" and "three-child" policies, the incidence of PAS is also rising rapidly, which is a tremendous problem faced by obstetrics in China.

Taking into account the high incidence and harmfulness of PAS, the academic community has conducted a lot of

explorations on the classification of PAS, hemostasis methods, and perinatal management, which improves the pregnancy outcome of PAS patients to a certain extent. However, the PAS classification is still more based on the depth of the placental villi invading the myometrium, and there are still few reports about the lateral extension of the placenta, especially the reports of cervical involvement [3]. In 2019, we retrospectively analyzed 96 cases of PAS in our hospital and found that PAS patients with cervical implantation had more bleeding volume during Cesarean section and higher probability of hysterectomy and bladder injury than those with lower uterine implantation, which indicated that it was necessary to focus on the placenta implantation in cervix. However, due to the lack of clinical experience and professional knowledge of medical imaging, PAS complicated with cervical implantation is easy to be misjudged by obstetricians, leading to excessive surgery, even unnecessary hysterectomy, or delay in hysterectomy, resulting in the increase of risk. Therefore, there is an urgent need to establish a high-performance risk early warning model, which is bound to change this unhealthy situation.

Artificial intelligence (AI) is developing rapidly in many areas, such as the medical industry. In the research of obstetrics and gynecology, AI has been applied to fetal monitoring [4], prediction of premature delivery [5], prediction of postpartum hemorrhage risk [6], prediction of vaginal trial delivery after Cesarean section [7], and forceps midwifery [8]. The disease risk warning model, a statistical model based on big data, is an outstanding performance of artificial intelligence in the medical field. Clinicians are able to identify high-risk groups, improve medical decision-making, and ultimately improve patient prognosis using the risk warning model. The present study retrospectively analyzed the clinical and imaging data of 166 pregnant women with PAS in the Maternal and Child Health Hospital Affiliated to Nanchang University from 2016 to 2020, screened the risk factors of PAS patients complicated with cervical implantation, and established a disease risk early warning model, so as to guide clinicians to evaluate the condition accurately and formulate a reasonable surgical treatment plan.

2. Data and Methods

2.1. Study Population. The data of pregnant women with PAS who were hospitalized and delivered in the Maternal and Child Health Hospital Affiliated to Nanchang University from January 1, 2016, to September 30, 2020, were collected retrospectively. All medical records were retrieved through the medical record management system. The retrieval strategies were placental implantation as the discharge code and Cesarean delivery as the operation code. The inclusion criteria included singleton pregnancy and clinical description diagnosis or postoperative pathological biopsy diagnosis of placental implantation. The exclusion criteria included pregnant women with primary coagulation dysfunction or other serious pregnancy complications and pregnant women with incomplete medical records. Finally, a total of 166 pregnant women were included in this study. This study was approved by the hospital ethics committee.

2.2. Diagnostic Criteria. The diagnosis of PAS was based on surgical and/or pathological evidence. Specifically, it was defined as a large number of placental villi invading myometrium, serosa, and even organs outside the uterus. The diagnostic criteria for PAS complicated with cervical implantation were the presence of the placenta that is attached to the internal cervix during Cesarean section and could not be peeled off by itself. Meanwhile, the cervix was found to be incomplete during manual dissection or the postoperative tissue biopsy indicated the incompleteness of cervix or the presence of villous tissue. The diagnostic criteria of PAS complicated with lower uterine segment implantation were that the placenta could not be peeled off by itself during Cesarean section. Simultaneously, only the lower uterine segment was invaded, the cervical mucosa was smooth and intact, and the villus or myometrium was missing in the myometrium of the uterus by tissue biopsy.

2.3. Methods of Medical Imaging Examination. All patients were examined by two fixed and experienced ultrasound or imaging doctors within one week before the Cesarean section. The model of the ultrasonic diagnostic apparatus was GE Voluson E8, and the probe frequency was set as 5–9 MHz. The ultrasonic examination methods were as follows. The patient was asked to empty the bladder before examination, and abdominal color Doppler ultrasound was performed to check the growth diameter of the fetus, the attachment position of placenta, and the relationship between placenta and the internal cervix. Subsequently, the patient was instructed to take the bladder lithotomy position. After wrapping the vaginal ultrasound probe with sterile condom, the probe was inserted into the fornix vaginae and was gently rotated to the sagittal section of the maternal uterus for measuring the length of the cervix, the width of the cervix, the blood flow and shape of the cervix, and the bladder line. For PAS patients with obesity, posterior placenta, or controversial ultrasound diagnosis, pelvic MRI was additionally performed to assist in the diagnosis. The model of the MRI diagnostic apparatus was SIEMENS superconducting 1.5 T magnetic resonance instrument. And the MRI examination methods were as follows. PAS patients adopt left decubitus position, and multichannel body phased array coils were used to image the placenta and cervix in cross-section, coronal, and sagittal planes.

2.4. Cesarean Section. Two chief physicians with corresponding qualifications in placental disease ward would perform this surgery. The main points of the Cesarean section were as follows. (1) Abdominal aortic balloon implantation was performed before Cesarean section for the patients with highly suspected cervical implantation. (2) Epidural anesthesia was the main anesthesia method and general anesthesia under tracheal intubation would be applied if necessary. (3) Bladder was separated and pushed down to the level of the intrauterine mouth before the fetus was delivered. (4) The placenta should be avoided as much as possible when you choose the uterine incision; if the placenta could not be avoided, we should choose the place where the

placenta was relatively thin. For patients with cervical implantation, the preset abdominal aortic balloon was filled at the same time as the fetus was delivered to block the blood supply of the uterus temporarily. (5) The uterus was dragged out and then the lower part of the uterus near the inner mouth was bundled up with the tourniquet. (6) When the placenta is implanted in the lower part of the cervix, we should strip out the placenta step by step with our hands and suture immediately. (7) The peeling surface of the placenta should be observed for about 15 minutes, and then the uterus was sutured if there was no active bleeding.

2.5. Estimation Method of Postpartum Bleeding Volume. According to the "Guidelines for Prevention and Treatment of Postpartum Hemorrhage" issued by the Obstetrics and Gynecology Committee of Chinese Medical Association, the volume of postpartum hemorrhage was evaluated by the combination of gauze weighing method and volumetric method. In the present study, negative pressure bottles and gauze were used to collect the bleeding during the perioperative period of emergency Cesarean section and the beginning of elective Cesarean section. The postoperative vaginal bleeding volume was calculated by the weighing method. The diagnostic criteria of postpartum hemorrhage, referring to the 2014 edition of clinical guidelines for diagnosis and treatment of postpartum hemorrhage, was that the bleeding volume within 24 hours after delivery was greater than 1000 ml.

2.6. Collection of Clinical Data. All clinical data was collected from our hospital's electronic medical record system, including basic information, the experience of the third trimester of pregnancy, imaging data, and delivery outcomes. The basic information included age, number of gravidities, times of induced abortions, and Cesarean sections. The third trimester of pregnancy mainly included the times of vaginal bleeding and the maximum volume of vaginal bleeding, and methods of volume and weighing were used to calculate the maximum vaginal bleeding volume. The imaging data included the attachment position of the main body of the placenta (anterior placenta or nonanterior placenta), the relationship between the placenta and the internal cervix (complete coverage or incomplete coverage), cervical length (distance between the internal and external cervical orifice), cervical width (left and right diameter of cervix), cervical blood flow (multiple anechoic areas of varying sizes in the cervix could be seen when cervical blood flow was abundant), cervical morphology, and bladder line (the high echo line between serosa and bladder cavity). The maternal outcome included gestational ages, postpartum bleeding volume, amount of blood transfusion, rate of hysterectomy, and bladder injury.

3. Statistical Analysis

IBM SPSS 24.0 software was used for data processing. The count data was expressed as frequency and rate, and the statistical analysis was performed with the chi-square test. The measurement data with a normal distribution was expressed as means and standard deviations, and the statistical analysis was performed using an independent sample t test. The measurement data with a skew distribution was expressed as median and quartile spacing, and the statistical analyses were performed using Mann–Whitney U test. Logistic regression was used to construct the risk prediction model of cervical implantation with PAS. ROC curve was used to evaluate the diagnostic ability of risk factors f cervical implantation with PAS. $P < 0.05$ was statistically significant.

4. Results

4.1. Distribution of Postpartum Bleeding Volume of the Two Groups. Among the 166 subjects, there were 49 cases of PAS complicated with cervical implantation and 117 cases of PAS complicated with lower uterine implantation. The postpartum hemorrhage rate of 68 patients with postpartum hemorrhage was 40.96% (68/166), of which the postpartum hemorrhage rate of cervical implantation group was 83.67% (41/49). The bleeding volume was mainly 1000 ml~2000 ml, and the postpartum bleeding volume of 4 patients was as high as 3000 ml. The postpartum hemorrhage rate in the lower uterine implantation group was 23.08% (27/117), and the bleeding volume group was mainly 500 ml~1000 ml. The chi-square test for trends showed that there was a linear trend between placental implantation position and postpartum bleeding volume. Details were shown in Table 1. The proportion of PAS complicated with cervical implantation increased with the increase of postpartum bleeding volume.

4.2. Maternal Outcomes in PAS Patients of the Two Groups. Comparing the maternal outcomes of PAS patients in the two groups, it was found that the amounts of bleeding volume, red blood cell transfusion, and plasma transfusion in Cesarean section in cervical implantation group were significantly higher than those in lower uterine implantation group. In addition, the hysterectomy rate and bladder injury rate in cervical implantation group were also significantly higher than those in lower uterine implantation group. See Table 2.

4.3. Univariate Analysis of PAS Complicated with Cervical Implantation. The independent sample t test, Mann-Whitney U test, and chi-square test were used to analyze the possible influencing factors of PAS complicated with cervical implantation. The results showed that the number of Cesarean sections, anterior placenta, complete placenta previa, bladder line interruption, cervical length, cervical width, abundant cervical blood flow, and cervical morphological disorder were all related to PAS complicated with cervical implantation. See Table 3.

4.4. Establishment and Validation of Logistic Regression Model for PAS Complicated with Cervical Implantation. The 8 independent variables obtained after univariate analysis were introduced into the logistic regression model, and the detailed results were shown in Table 4. The effects of eight variables in the prediction model were analyzed. The results showed that the cervical width (X_1), abundant cervical blood

Table 1: Distribution of postpartum bleeding volume of the two groups.

Postpartum bleeding volume (ml)	Cervical implantation group (case/rate)	Implantation group (case/rate)
<500 ($n = 35$)	2 (5.70%)	33 (94.30%)
500~($n = 63$)	6 (9.50%)	57 (90.50%)
1000~($n = 48$)	23 (47.90%)	25 (52.10%)
2000~($n = 16$)	14 (87.50%)	2 (12.50%
3000 ($n = 4$)	4 (100%)	0 (0)
Linear-by-linear association	57.047	
P value	<0.001	
Method	Monte Carlo	

Table 2: Maternal outcomes of the two groups.

	Cervical implantation group ($n = 49$)	Lower uterine implantation group ($n = 117$)	P value	Method
Bleeding volume (ml)	1800 (1000, 2100)	600 (400, 900)	<0.001	Mann-Whitney
Red blood cell transfusion (ml)	900 (500, 1450)	0 (0, 450)	<0.001	Mann-Whitney
Plasma transfusion (ml)	400 (0, 700)	0 (0, 100)	<0.001	Mann-Whitney
Hysterectomy (case/rate)	6 (12.25%)	0 (0)	0.001	Continuity correction
Bladder injury (case/rate)	6 (12.25%)	1 (0.86%)	0.001	Continuity correction

Data are presented as median (interquartile spacing).

Table 3: Univariate analysis of PAS complicated with cervical implantation.

	Cervical implantation group	Lower uterine implantation group	P value	Method
Age (years)	32 (30, 35)	30 (28, 34)	0.071	Mann–Whitney
Gravidity (times)	4 (3, 6)	4 (3, 5)	0.404	Mann–Whitney
Induced abortion (times)	1 (1, 3)	1 (1, 2)	0.790	Mann–Whitney
Cesarean section (times)	1 (1, 2)	1 (1, 2)	0.017	Mann–Whitney
Vaginal bleeding (times)	1 (0, 1.50)	1 (0, 2)	0.486	Mann–Whitney
Maximum volume of vaginal bleeding (ml)	10 (0, 90)	10 (0, 120)	0.767	Mann–Whitney
Anterior placenta (case/rate)	41 (83.70%)	64 (54.70%)	<0.001	Pearson
Complete placenta previa (case/rate)	48 (98%)	95 (81.20%)	0.004	Pearson
Cervical length (mm)	28.06 ± 6.37	33.67 ± 7.49	<0.001	Independent sample t
Cervical width (mm)	36.96 ± 6.22	28.25 ± 3.08	<0.001	Independent sample t
Abundant cervical blood flow (case/rate)	37 (75.50%)	3 (2.60%)	<0.001	Pearson
Morphological disorder of cervix (case/rate)	29 (59.20%)	5 (4.30%)	<0.001	Pearson
Bladder line interruption (case/rate)	38 (77.60%)	11 (9.40%)	<0.001	Pearson

Data are presented as median (interquartile spacing) or mean ± standard deviation.

flow (X_2), and bladder line interruption (X_3) contributed significantly to the regression model, which were independent risk factors of PAS complicated with cervical implantation, while the other five independent variables had little effect. Therefore, the prediction model was established as follows:

$$\hat{P} = \frac{exp\left(-16.584 + 0.352X_1 + 5.036X_2 + 2.514X_3\right)}{1 + exp\left(-16.584 + 0.352X_1 + 5.036X_2 + 2.514X_3\right)}. \quad (1)$$

The Hosmer–Lemeshow goodness-of-fit test showed 5.984 of χ^2 and 0.649 of P value, indicating that the model had a good fit. The sensitivity of the prediction model was 91.84%, the specificity was 98.29%, the missed diagnosis rate was 8.16%, and the misdiagnosis rate was 1.71%.

4.5. Joint Diagnosis of PAS Complicated with Cervical Implantation by Multiple Independent Risk Factors. In this study, the logistic regression analysis showed that cervical width, abundant cervical blood flow, and bladder line interruption were independent risk factors of PAS with cervical implantation. When the cervical width was used as the diagnostic index alone, the optimal diagnostic threshold of cervical width was 33.5 mm and AUC was 0.854. When the abundant cervical blood flow and bladder line interruption

Table 4: Logistic regression analysis of PAS complicated with cervical implantation.

	B	SE	Wald	*P*	OR	95% CI
Cesarean section	−0.749	0.628	1.420	0.233	0.473	(0.138, 1.621)
Anterior placenta	−0.405	0.900	0.203	0.653	0.667	(0.114, 3.893)
Central placenta previa	−4.182	2.734	2.340	0.126	0.015	(0.000, 3.241)
Cervical length	−0.033	0.066	0.253	0.615	0.967	(0.851, 1.101)
Cervical width	0.352	0.111	9.987	0.002	1.421	(1.143, 1.767)
Abundant cervical blood flow	5.036	1.281	15.462	<0.001	153.851	(12.502, 1893.379)
Morphological disorder of cervix	1.472	0.970	2.303	0.129	4.356	(0.651, 29.136)
Bladder line interruption	2.514	0.940	7.151	0.007	12.353	(1.957, 77.976)
Constant	−16.584	5.423	9.352	0.002		

were taken as diagnostic indexes, the corresponding AUC were 0.865 and 0.841, respectively. The joint diagnosis of these three independent risk factors showed 0.886 of AUC. Details were shown in Figure 1.

5. Discussion

Currently, the etiology of PAS remains controversial, and the mainstream etiology theory holds that its onset may be attributed to the secondary local tissue hypoxia during the process of uterine scar formation, resulting in abnormal vascularization, which further leads to the defect of uterine decidua and abnormal invasion of trophoblast cells into uterine myometrium and pelvic organs. The cervix, a special part of the uterus, is normally located in the deep position of the female pelvic cavity, mainly composed of connective tissue, with few smooth muscle fibers, blood vessels, and elastic fibers, and mainly receiving blood supply from the descending branch of uterine artery and part of vaginal artery. Once the placenta invades the cervix, the blood flow of the cervix increases, and the difficulties such as poor contraction of myometrium, wide bleeding wound, fast bleeding speed, and difficulty in stopping bleeding easily occur in the process of placental abruption [9], which seriously perplex the majority of obstetricians.

In 2020, Liu analyzed the clinical data of 105 cases of invasive placenta previa and found that 7 patients with cervical implantation had undergone hysterectomy due to severe bleeding [10]. In this study, 49 cases of PAS had cervical implantation, of which 41 cases had postpartum hemorrhage, 6 cases had hysterectomy due to intractable postpartum hemorrhage, and 6 cases had bladder injury due to placenta penetrating the cervix and invading the bladder. It was not unexpected to find out that the incidence of these three adverse outcomes in cervical implantation group was significantly higher than that in lower uterine implantation group by comparing the postpartum hemorrhage rate, hysterectomy rate, and bladder injury rate between cervical implantation group and lower uterine implantation group. As an obstetrician, it is of great clinical significance to determine whether PAS is complicated with cervical implantation before operation, so as to improve the safety of Cesarean section and improve the maternal outcome. Therefore, the early identification of the risk factors of PAS complicated with cervical implantation and the construction of a risk warning model are particularly essential.

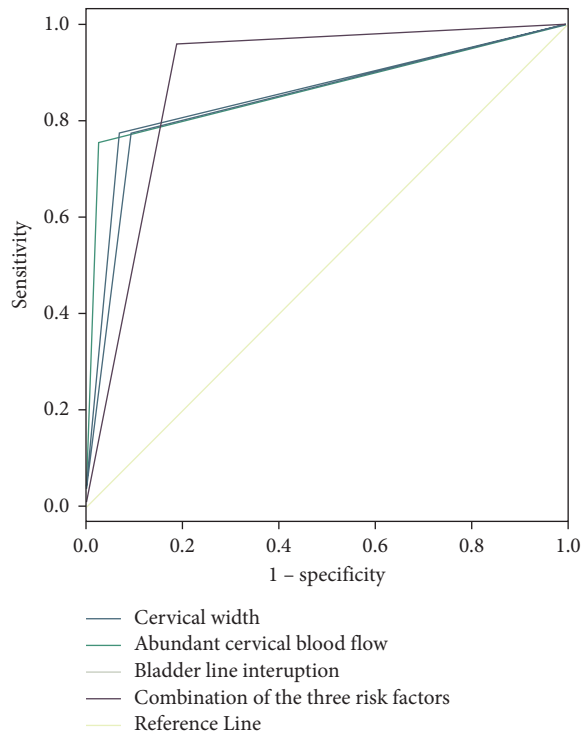


Figure 1: Risk prediction model and ROC curve of each risk factor.

Clinical indicators and medical imaging data are irreplaceable for obstetricians to assess the condition of PAS patients. The clinical indicators mainly include the patient's demographic characteristics, past medical history, and pregnancy history. The subjects included in this study were pregnant women with placenta previa complicated with placenta implantation. The clinical indicators included the clinical data of age, times of pregnancies, induced abortions, Cesarean sections, vaginal bleeding in the third trimester, and maximum volume of vaginal bleeding. Univariate analysis found that there was a correlation between the number of Cesarean sections and PAS complicated with cervical implantation, which was consistent with the current global PAS consensus, indicating that the incidence rate of PAS increased with the increase of Cesarean delivery [11–13]. Obstetricians are able to make a preliminary assessment of the disease through comprehensive analysis of various clinical indicators and decide on the supplementary examination that needs to be improved in the followup. At

present, multimodal ultrasound imaging has been recommended by the diagnosis and treatment guidelines of PAS in various countries to accurately assess the high-risk women with PAS [14]. In the present study, color Doppler ultrasound was used as the preferred auxiliary examination, including transabdominal ultrasound and transvaginal ultrasound [15]. For patients with obesity or placenta in posterior wall, we performed pelvic MRI [16]. In addition to the common ultrasound signs of PAS, we paid special attention to cervix-related images in the included image data, including the length, width, blood flow, and morphology of the cervix. A univariate analysis of 7 imaging signs showed that these imaging signs were all related to PAS complicated with cervical implantation.

During the past years, the observation time window of cervical morphology was mainly in the second trimester of pregnancy which aimed to predict the incidence of preterm delivery. In recent years, based on the idea that changes of cervical morphology could reflect the changes of cervical load and surrounding structure, the observation indexes of cervical morphology have been gradually brought into the scope of imaging examination of placenta previa with placenta accrete. Polat et al. [9] thought that cervical shortening seemed to be a warning sign of placenta previa with penetration. Our study showed that the cervical length of PAS patients complicated with cervical implantation was significantly shorter than that of lower uterine implantation, which was consistent with Polat M's study. However, when the number of Cesarean sections and 7 imaging signs were used as independent variables for logistic regression analysis, the results showed that cervical width, abundant cervical blood flow, and bladder line interruption were independent risk factors of PAS complicated with cervical implantation. The optimal threshold value of cervical width was 33.5 mm. When other independent variables were fixed, the probability of cervical implantation in PAS increased by 0.421 times for every 1 mm increase in cervical width. When imaging examination showed abundant cervical blood flow, the probability of cervical implantation in PAS increased 152.851 times. In addition, the bladder is the organ closest to the lower part of the uterus and the cervix in the third trimester. The bladder line can reflect the relationship between bladder, lower part of the uterus, and cervix. If ultrasound or MRI indicates that the bladder line is interrupted, it means that the boundary between the bladder and the uterus is not clear. In this study, when the imaging examination showed that there was bladder line interruption, the probability of cervical implantation in PAS increased by 11.353 times. ROC analysis showed that the corresponding AUC were 0.854, 0.865, and 0.841, respectively, when cervical width, cervical blood flow, and bladder line interruption were used alone for diagnosis of PAS with cervical implantation. While the AUC was 0.886 when the three indexes were used for joint diagnosis, which indicated that the combined detection had good diagnostic value for PAS with cervical implantation.

There was no gainsaying the fact that, due to the nature of a retrospective single-center cohort study, the present study excluded some cases of missing data or insufficient imaging data, cases of twins, and pregnant women with multiple fetuses, which may have a certain impact on the results of the statistical analysis. Concurrently, the lack of data on the length and width of cervix in 166 patients with placental implantation before pregnancy or early pregnancy makes this study unable to accurately and dynamically evaluate the changes of cervix during pregnancy. More reliable models should be established with the help of the combined analysis of multicenter data in future researches.

6. Conclusion

PAS complicated with cervical implantation was extremely severe, which is often accompanied by adverse pregnancy outcomes such as postpartum hemorrhage, hysterectomy, and bladder injury. Preoperative image of Cesarean section has certain predictive value in predicting PAS with cervical implantation. Establishing the risk prediction model of PAS complicated with cervical implantation by logistic regression analysis before operation has positive significance in guiding obstetricians to formulate reasonable treatment plan.

Data Availability

The labeled dataset during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare no potential conflicts of interest.

Authors' Contributions

S. Xin and H. Wan contributed equally to this work.

Acknowledgments

This work was supported by the Technology Plan of Jiangxi Provincial Health Commission (nos. 20203624 and 20203614) and the Science and Technology Project of Jiangxi Province (no. 20212BAB216065).

References

[1] Y. Berhan and T. Urgie, "A literature review of placenta accreta spectrum disorder: the place of expectant management in Ethiopian setup," *Ethiopian journal of health sciences*, vol. 30, no. 2, pp. 277–292, 2020.

[2] E. Jauniaux, A. M. Hussein, K. A. Fox, and S. L. Collins, "New evidence-based diagnostic and management strategies for placenta accreta spectrum disorders," *Best Practice & Research Clinical Obstetrics & Gynaecology*, vol. 61, pp. 75–88, 2019.

[3] E. Jauniaux, S. L. Collins, D. Jurkovic, and G. J. Burton, "Accreta placentation: a systematic review of prenatal ultrasound imaging and grading of villous invasiveness," *American Journal of Obstetrics and Gynecology*, vol. 215, no. 6, pp. 712–721, 2016.

[4] J. Balayla and G. Shrem, "Use of artificial intelligence (AI) in the interpretation of intrapartum fetal heart rate (FHR) tracings: a systematic review and meta-analysis," *Archives of Gynecology and Obstetrics*, vol. 300, no. 1, pp. 7–14, 2019.

[5] R. O. Bahado-Singh, J. Sonek, D. McKenna et al., "Artificial intelligence and amniotic fluid multiomics: prediction of perinatal outcome in asymptomatic women with short cervix," *Ultrasound in Obstetrics and Gynecology*, vol. 54, no. 1, pp. 110–118, 2019.

[6] Y. Miyagi, K. Tada, I. Yasuhi et al., "New method for determining fibrinogen and FDP threshold criteria by artificial intelligence in cases of massive hemorrhage during delivery," *Journal of Obstetrics and Gynaecology Research*, vol. 46, no. 2, pp. 256–265, 2020.

[7] M. Lipschuetz, J. Guedalia, A. Rottenstreich et al., "Prediction of vaginal birth after cesarean deliveries using machine learning," *American Journal of Obstetrics and Gynecology*, vol. 222, no. 6, pp. 613–e12, 2020.

[8] S. Xin, Z. Wang, H. Lai et al., "Clinical effects of form-based management of forceps delivery under intelligent medical model," *Journal of Healthcare Engineering*, vol. 2021, p. 9947255, 2021.

[9] M. Polat, I. Kahramanoglu, T. Senol, E. Ozkaya, and A. Karateke, "Shorter the cervix, more difficult the placenta percreta operations," *Journal of Maternal-Fetal and Neonatal Medicine: The Official Journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstetricians*, vol. 29, no. 14, pp. 2327–2331, 2016.

[10] W. Liu, X. Chen, C. Sun, X. Wei, G. Wang, and R. Shan, "Morphological evaluation of cervix using MRI at 32 to 36 weeks of gestation," *Medicine*, vol. 97, no. 49, p. e13375, 2018.

[11] E. Jauniaux, F. Chantraine, R. M. Silver, and J. Langhoff-Roos, "FIGO consensus guidelines on placenta accreta spectrum disorders: e," *International Journal of Gynecology & Obstetrics*, vol. 140, no. 3, pp. 265–273, 2018.

[12] R. M. Silver, M. B. Landon, D. J. Rouse et al., "Maternal morbidity associated with multiple repeat cesarean deliveries," *Obstetrics & Gynecology*, vol. 107, no. 6, pp. 1226–1232, 2006.

[13] E. Jauniaux, R. M. Silver, and S. Matsubara, "The new world of placenta accreta spectrum disorders," *International Journal of Gynecology & Obstetrics*, vol. 140, no. 3, pp. 259-260, 2018.

[14] E. Jauniaux, J. C. Kingdom, and R. M. Silver, "A comparison of recent guidelines in the diagnosis and management of placenta accreta spectrum disorders," *Best Practice & Research Clinical Obstetrics & Gynaecology*, vol. 72, pp. 102–116, 2021.

[15] E. Jauniaux and D. Ayres-de-Campos, "FIGO consensus guidelines on placenta accreta spectrum disorders: i," *International Journal of Gynecology & Obstetrics*, vol. 140, no. 3, pp. 261–264, 2018.

[16] E. M. Berkley and A. Z. Abuhamad, "Prenatal diagnosis of placenta accreta," *Journal of Ultrasound in Medicine*, vol. 32, no. 8, pp. 1345–1350, 2013.

Hindawi
Journal of Healthcare Engineering
Volume 2022, Article ID 4277436, 14 pages
https://doi.org/10.1155/2022/4277436

Research Article

AI-DRIVEN Novel Approach for Liver Cancer Screening and Prediction Using Cascaded Fully Convolutional Neural Network

Piyush Kumar Shukla,[1] Mohammed Zakariah,[2] Wesam Atef Hatamleh,[3] Hussam Tarazi,[4] and Basant Tiwari[5]

[1]*Computer Science & Engineering Department, University Institute of Technology, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal 462033, India*
[2]*College of Computer and Information Sciences, King Saud University, P.O. Box 51178, Riyadh 11543, Saudi Arabia*
[3]*Department of Computer Science, College of Computer and Information Sciences, King Saud University, P.O. Box 51178, Riyadh 11543, Saudi Arabia*
[4]*Department of Computer Science and Informatics, School of Engineering and Computer Science, Oakland University, Rochester Hills MI USA 318 Meadow Brook rd, Rochester, MI 48309, USA*
[5]*Department of Information Technology, Hawassa University, Institute of Technology, Hawassa, Ethiopia*

Correspondence should be addressed to Basant Tiwari; basanttiw@hu.edu.et

Received 18 October 2021; Revised 18 December 2021; Accepted 5 January 2022; Published 1 February 2022

Academic Editor: Xingwang Li

In experimental analysis and computer-aided design sustain scheme, segmentation of cell liver and hepatic lesions by an automated method is a significant step for studying the biomarkers characteristics in experimental analysis and computer-aided design sustain scheme. Patient to patient, the change in lesion type is dependent on the size, imaging equipment (such as the setting dissimilarity approach), and timing of the lesion, all of which are different. With practical approaches, it is difficult to determine the stages of liver cancer based on the segmentation of lesion patterns. Based on the training accuracy rate, the present algorithm confronts a number of obstacles in some domains. The suggested work proposes a system for automatically detecting liver tumours and lesions in magnetic resonance imaging of the abdomen pictures by using 3D affine invariant and shape parameterization approaches, as well as the results of this study. This point-to-point parameterization addresses the frequent issues associated with concave surfaces by establishing a standard model level for the organ's surface throughout the modelling process. Initially, the geodesic active contour analysis approach is used to separate the liver area from the rest of the body. The proposal is as follows: It is possible to minimise the error rate during the training operations, which are carried out using Cascaded Fully Convolutional Neural Networks (CFCNs) using the input of the segmented tumour area. Liver segmentation may help to reduce the error rate during the training procedures. The stage analysis of the data sets, which are comprised of training and testing pictures, is used to get the findings and validate their validity. The accuracy attained by the Cascaded Fully Convolutional Neural Network (CFCN) for the liver tumour analysis is 94.21 percent, with a calculation time of less than 90 seconds per volume for the liver tumour analysis. The results of the trials show that the total accuracy rate of the training and testing procedure is 93.85 percent in the various volumes of 3DIRCAD datasets tested.

1. Introduction

1.1. Stages of Liver Cancer. In the biomarker selection of illness, the anatomical study of the liver and divisible lesions on magnetic resonance imaging are used in the selection of disease biomarkers. The diagnostic phases, succession initial and secondary hepatic tumours analysis are used [1]. The majority of initial tumours in different organs, such as the liver, colon area, and pancreatic region, commonly spread to the smaller structures in the organ. Consequently, frequent examination of the liver and its lesions is required in order to determine the main stage of a cancerous tumour. In addition

to hepatocellular carcinoma illness, another major cause for infection of the liver area exists. In the liver area, this illness is referred to as a primary tumour disease, and it is one of the sixth most prevalent cancer diseases in the world, as well as the third most common cause of death for cancer patients globally [2–5]. Hepatocellular carcinoma is a kind of cancer that is genetic and molecular in nature, and it is most often caused by a chronically injured liver.

Various hepatocellular carcinoma illnesses have been identified and grouped into distinct groups based on their clinical presentation [6]. The progressive development of hepatocellular carcinoma is dependent on changes in tissue architecture and variations in vascular supply. Changes in the architecture of the tissue have been shown to accelerate the formation of additional tissue in the liver, which has been discovered via the use of medical imaging [3]. This procedure is dependent on the tumor cells' architectures and shapes, as well as their sizes. In clinical diagnosis, CT and MRI scans are utilised to examine liver cancer, with physical or semi-manual segmentation methods being employed in the process. These approaches are manual, highly operational, subjective, and time-consuming, and they need greater effort. Through the automated function, it is possible to decrease the amount of time spent on invention and radiologist improvement in computer-aided techniques, and to build unique segmentation methods. Automatic segmentation was performed on the combined liver and lesion area picture [7] to determine the extent of the lesion. The uneven segmentation in low contrast pictures between the liver and the lesion site has proven to be a significant hurdle for the researchers to overcome. When comparing hyper and hypo tumors, the contrast levels may be different, and the aberrant tissue development in the lesion may be different in different sizes and numbers of the lesion [8]. It is not possible to segment the liver area using the intensity-based technique because of the intricacy of the contrast variations seen across several testing instances. Cancer cells may have a variety of shapes, which reduces the efficacy of computational approaches that segment cancer cells. The suggested technique differentiates between cancer stages based on the structure of the tumor and the form of the lesion.

1.2. Shape Parameterization. In the image-based registration approach, the shape analysis is the most important component in segmenting the tumor area from the lesion region [4] and determining the location of the tumor. The memory of form, size, and tumor structure with respect to the metrics and landmarks of the liver cells is part of the image-based registration approach. Image-based registration technique This unique parameterization [9], which is based on the difference between the two objects, integrates the shape descriptor and the two-point inconsistencies in the picture in a single step. It has been shown that the prior approach for surface analysis may be used in a medical setting [5]. The parameterization of human organs is a difficult job to do throughout the segmentation phase. It is the sophisticated segmentation in medical imaging that is characterized by the parameterization of star-shaped objects in the abdominal region [10]. To lessen the deficiency impact, the suggested technique employs injective mapping in 2D surfaces, therefore removing the image's various areas [11]. The physicians, on behalf of the liver segmentation, validated the matching spots that would be used to include the statistical ship models throughout the training phase of the parameterization process.

1.3. Liver Segmentation. In recent years, researchers have devised a promising and new approach for detecting cancer and metastases. The most recent advancements in building and design have created new chances for metastasis to develop. Aside from that, dynamic bimolecular settings for various tumors are being researched. Types. Because of this, we may conclude that there is a causal relationship between gene expression levels in different tissues. It is possible to have a better understanding of how cancer develops early on biomolecular networks that have been connected to both normal and pathological processes by investigating various types of cancer. Cancerous states are a kind of cancer in and of itself. According to Ling et al. (2014), this hypothesis has been explored and he concludes that he has researched if there is a link between the mRNA terminologies of three distinct genes.

Following a random selection of the cancer-related genes PIK3C3, PIM3, and PTEN, the researchers discovered that the cancer had progressed. Since then, these coefficients have been tested in the area of cancer research. Diagnosis. While the patient was sick, the following observations were taken, and a decision was formed on how to treat him: A substantial correlation of 0.68 r 1.0 was found between the variables, indicating that the variables were related. PIM3 and PIK3C3 in breast cancer, PIM3 and PTEN in breast, kidney, and ovarian cancer, and PIM3 and PTEN in prostate cancer Malignancies of the liver and thyroid, as well as cancers of the breast and ovary, have been linked to PIK3C3 and PTEN mutations in the past. There is an assumption that the connections for early cancer detection are necessary in order to integrate the gene expression profiles of cancer networks to the clinical data that is already available. Biomarkers include things like cancer antigens and other such things. About ten to fifteen percent of all human malignancies are caused by cancers. Viruses are also responsible for certain cancers. A technique known as massively parallel sequencing has been found to be successful in both malignancies and normal tissue for the discovery of new viruses and the interactions between them.

1.4. Analyses of Hepatic Tumors (1.4). According to MICCAI, the problem of differentiation was first raised in 2008, at a time period that included the segmentation of liver tumors [12, 13]. It researches and develops different disease segmentation strategies as well as contrast improvement approaches for tumor segmentation from healthy hepatic parenchyma utilizing computed tomography images. Contained Participants were given an introduction to data and measuring procedures. Tumor segmentations were tested using five semi-automated and four automatic methods. The

approaches for the liver differentiation competition were estimated using equivalent metrics. To separate tumors [14], Standard Graph Cutting techniques and the watershed algorithm were utilised suitably, much as they did for liver segmentation methods, to provide the most accurate measuring tool. A similar number of semi-automatic approaches are used; the most frequent are adaptive thresholds and morphologies, voxel identification and dissemination, and pixel identification, among others. In order to segment tumors, neural network technology and an ad boost taught community discriminating by artificial intelligence and picture recognition are the most effective approaches [15].

1.5. Our Approach. We will apply new automated technologies for the segmented liver, which will continuously improve contrast and imaging artefact removal while reducing the amount of time required. The resilient parameter of 3-D surfaces is presented for use in the segmentation procedure among abdominal organ pictures in order to increase the contrast between the two images. On the surface of objects, the 3-dimensional plane is represented in the form of the x, y, and z-axis when a closed planar curve is drawn on the surface. This representation of the space eliminates the frequent issues associated with the surface parameterization of concave objects. In order to eliminate noise type descriptors throughout the segmentation process, we use rotational and resilient approaches. A shape-driven geodesic active contour is used to improve liver segmentation after the first segmentation has been identified [16]. It is necessary to detect and treat hepatic tumours at an early stage in order to reduce the risk of mortality in a given individual. The features of tumour candidates are retrieved, and the support vector machine Algorithm is used to classify the candidates. It has been determined that the suggested segmentation approach outperforms and can be compared to current algorithms on a number of datasets with varying age limitations. In the case of hepatic tumour imaging, the segmentation of the liver and the computation of tumour measurement are difficult. Patients with abnormal livers may be traced down and identified using the automated liver cancer analysis method, even if their photos have poor image quality or are incorrectly labelled.

Because of the limited amount of medical image data that is currently accessible, as well as the limits imposed by GPUs, the exploitation of 3D data may result in overfitting challenges in certain cases. This research proposes an enhanced VNet and a 2.5-dimensional convolutional neural network VNet WGAN to acquire context information from 3D data in order to achieve end-to-end segmentation of liver images. The enhanced VNet and the 2.5-dimensional convolutional neural network VNet WGAN are used to achieve end-to-end segmentation of liver images. Among their key tasks are the following:

In this step, two convolution kernels are used in series with the input being the stack of slices and their upper and lower adjacent slices, and the output being the segmentation map corresponding to the central slice to fully extract the intralayer and interlayer information that will be used in the 3D liver model.

(1) The input being the stack of slices and their upper and lower adjacent slices, and the output being the segmentation map corresponding to the central slice Despite the fact that segmentation accuracy remains high, geographic considerations may be able to assist reduce the amount of memory used and the amount of computation required.

(2) In order to make full use of the network's high-level and low-level characteristics, a chain residual pooling module is added to the VNet network's long-skip link structure in order to maximise the use of both high-level and low-level features. This enables for the accumulation of more detailed semantic information as well as an increase in the accuracy of liver segmentation by a large margin.

(3) Incorporate the boundary loss function into the basic WGAN generator network to make up for the lack of attention paid to the marginal pixel accuracy of the Dice loss function in the preceding step by including the boundary loss function. By including the composite loss function of boundary and Dice weighted fusion into the equation, the segmentation ability of the model is enhanced from the region and the boundary, respectively.

The ability to take random inputs and generate the appropriate output, as well as execute efficient inference and learning processes, is shown by fully convolutional networks that are cascaded. Adversity is a part of life. Through the use of this approach, the function is evaluated over the whole image frame. Patch-based approaches, as well as segmentation objects, are also utilised in this application. Instead of processing patches, the network processes entire pictures, reducing the amount of time spent on the network and the requirement to choose fertile areas in order to minimise the amount of repetitive reproduction estimate when patches overlap, resulting in an increase in the size of the final picture. The House of Representatives passed a resolution. Furthermore, many scales are integrated by connecting them together in ways that combine the most recent detection with lower layers with higher resolution. These measurements are made possible by combining numerous scales. This kind of fusion may be created in a number of shapes and sizes. This procedure generates a heat map of the lesion, which may then be used to diagnose it if necessary.

The following are some of the significant contributions made by this paper:

(i) It is our intention to offer the Cascaded Fully Convolutional Neural Network, which will be utilised for the detection and segmentation of liver cancer.

(ii) It is being developed with the assistance of a deep learning system that is effective in segmentation and classification. Tumors of the liver are categorised based on where they are found on the body.

(iii) The experimental results reveal that the proposed HFCNN is successful in that it makes use of the dataset to achieve high overall performance.

It is also anticipated that the suggested approach would assist the individual with the pace at which tumour cells develop, which will aid in the early identification and diagnosis of cancer. The following sections are included in the paper that was submitted. Section 2 contains a list of comparable works that are discussed in the context of the background study, and Section 3 contains the methodology for the proposed work. Section 4 discusses the experimental examination, and Section 5 summarises the findings and discusses future research opportunities.

2. Related Works

In CT imaging of the liver and liver lesions, there are numerous techniques for segmenting the organ that have been developed in both interactive and automated approaches. [17] Two benchmarks were conducted on liver and liver lesions segmentation at the MICCAI 2007 and 2008 Sessions, which were both published in [18]. The statistical model forms were the focus of the concerns discussed throughout the workshop. In addition, the workshop focused on grey levels and lesions texture analysis [18], which were also discussed. Otsu segmentation is a method that has lately become popular for graph cutting and level setting in pictures of liver cancer. However, because of the rise in velocity and intensity level, as well as the poor contrast in CT data, these approaches are not routinely employed in clinical settings. Interactive approaches are continuously being developed to address these flaws and strengthen their defences against future attacks. Target identification, classification, and segmentation are among the computer-vision problems that the academic community is becoming increasingly concerned with, thanks to algorithms such as cutting-edge techniques, Deep Convolutional Neural Networks (CNN) [19], which are used to do these tasks. Above all, CNN techniques have been shown to be the most user-friendly and most novel methodology for the segmentation of liver cancer in CT images as well as the segmentation of lesions in CT images, and they are now the most widely used.

According to Rong Zhu et al. [20], an image processing filter application of an improved anisotropic diffusion was developed, showing that anisotropic diffusion filters are the most frequent strategy for noise reduction. This study describes a more efficient approach for the anisotropic diffusion filter, which may be used to remove salt and pepper noise from photos. Ravi S, et al. [21] developed Morphological Operations for Image Processing, Understanding, and Applications, which they put into practise. The purpose is to remove any defects that may exist within the picture structures. Wassim Abdulrahman and colleagues [22] The term "Segmentation of Liver Tumors Using Image Processing" refers to the process of distinguishing specific parts of the liver in abdominal CT scans. When it comes to retrieving the tumor's location from CT scans, a new method has been developed. Amit Verma and colleagues, [23] "Techniques for Detecting Tumors Using Digital Imagery The background analysis in the segmentation of tumour cells is provided by the survey." If you compare the performance of this method to other existing methodologies, it is the most effective at finding and categorising cancers. According to Jinshan Tang et al. [24], an Adjustable Anisotropic Noise Reduction filter in MR images was developed, and it was recommended that an adaptive threshold range be used in the stepped-forward anisotropic diffusion filter. A transparent diffusion with an anisotropic probability-pushed memory system is proposed to tackle the over filtering problem by selecting a tissue and an overall metaphysical impact from a large number of possible options. The proposed technique has been tested in real MR pictures, and the results have been outstanding.

Alireza is an anisotropic diffusion filter that is used to cancel noise in the input picture throughout the processing steps. R. Lin and E. K. Wong [25] developed "Morphological operations on quadrants represented by images," which included a series of guidelines for performing direct morphological operations on quadrates represented by images as well as creating dilated and eroded snap images representing quadrants that were based on quadrates. Ruchika Chandel et al. [26] defined the segmentation algorithms and technique used for illustration of filter in embodiments and smoothing, as well as the smoothing algorithms and technique "Image Filtering Algorithms and Techniques Image smoothing, also known as image smoothing algorithms and techniques, is one of the most significant image dispensation techniques that is widely used. According to N. Howard and colleagues [27]," "a novel completely automated liver and tumour fragmentation system with a morphological operation" was developed for a numerical hepatocellular carcinoma detection method that was both high-sensitive and low-specific in its imaging. The present system came to the conclusion that segmentation using 2-dimensional photos is less accurate and requires more time to analyse than 3-dimensional images.

Prior to anything else, it's important to segment the liver so that the tumour on the CT image may be appropriately segmented. The segmentation of a tumour, and much more so the segmentation of a tumour in combination with the liver, is substantially more complex than it seems at first glance. General practitioners (GPs) will have a tough time visually distinguishing the liver and tumour from other undesired cells and nearby organs if they do not have specific training in this area. When a CT picture has low contrast, a dynamic size, non-uniform intensity, and an assortment of artefacts, segmenting the liver and tumour is considerably more challenging, even for an experienced radiologist or doctor. However, despite the fact that segmentation conducted by professionals is accurate, it requires a large investment of time and effort. Apart from that, specialists in liver cancer who are capable of executing exact and fine segmentation are rare to find, and they are especially inaccessible to the people of impoverished countries, where the issue of liver cancer is more frequent. A more exact and effective algorithm for tumour segmentation, as well as algorithms for assessing tumour size, shape, and location,

are required as a result of the presence of these problems. A range of semi-automatic and manual procedures have been developed in order to segment the tumour in the liver and determine its location. Given that each of these systems is dependent on human interaction, they are all susceptible to user error, individual bias in feature selection, and time lag. A totally automated segmentation technique, capable of segmenting both the liver and the tumour in a single run, is necessary in order to do this. In this way, a doctor or radiologist may reduce reading time, increase detection sensitivity, enhance diagnostic accuracy, and discover malignancies early in the process without interfering with the patient's health. Following extensive testing, these technologies may also be used in instances where there is a paucity of competence in liver imaging. In recent years, researchers have concentrated their efforts on creating a wholly automated system that can generate accurate and timely forecasts of liver tumours while saving a large amount of time and effort on the part of the researcher. The benefit of entirely automated approaches is that they evolve over time as a consequence of their output as well as the absorption of diverse conditions and inputs into the system as the system matures. A large number of research that have lately been published provide credence to this notion.

When using 3-dimensional pictures with the geodesic active contour approach, the suggested methodology boosts the accuracy rate while simultaneously decreasing the segmentation processing time, hence eliminating these flaws.

3. Methods and Materials

3.1. Overview of Our Proposed Segmentation Processing. 380 patients contributed a total of 2012 CT images, which was collected from 398 individuals. Three hundred and thirty-three patients with Hepatocellular Carcinoma in Adults were found, resulting in a total of 591 CT pictures; three hundred and twenty-five patients with Hepatocellular Carcinoma in Children were identified, generating a total of 1421 CT images In order to establish the final diagnosis of these photographs, and in the absence of surgical intervention, the results of the lesions were utilised to establish the facts, so enabling the data to be regarded as reliable. A qualified physician additionally changed the window width and window level of the CT scans shown above to guarantee that the cancer tissue could be clearly seen in each image. Using these modifications, the cancer tissue was clearly seen in each image. The Digital Imaging and Communications in Medicine (DICOM) data was utilised to construct the final picture after it had been normalised to a grayscale image with a grayscale value of 0–255 according to the appropriate window width and window level. GIF files include the jpg extension, which stands for grayscale picture format. Using a medical professional's hands, the shape of the liver region was created in the CT image. The intended work will be divided into three main phases. The first stage is concerned with the preprocessing and segmentation of data. The graph cut approach is used in the second phase to gather features based on the segmented area of lesions that have been identified. In the final third stage, two cascaded fully convolutional neural networks with Training and Testing pictures are used to solve the problem.

In Figure 1, we can see the suggested workflow, which is made up of all of the various processing units. For the purpose of performing early-stage detection of liver cancer, the training and testing stages are carried out.

3.2. Shape Analysis and Surface Parameterization. As a S shape function, a 3D equivalent of a curve characteristic was utilised to equate closed planar curves in order to discover structural discrepancies in the data. S is the cross section area of the interior of the turn and "seed" at the given point p on a planar curve, in further detail (a sphere centred at p). If the instantaneous parameterization of the two curves is adequate, a S curve C may be used for both regional and global comparisons of the two curves, assuming that the two curves are sufficiently parameterized (at any corresponding position on two matched curves). Var is the volume of the C-intersection, and the Radius r radius seed sphere matches the size of the C-intersection. The influx of digital goods served as a consoling element in the situation. The mode descriptor is invariant with respect to architecture and resilient to noise.

Our method makes point-to-point comparisons across numerous surfaces by using the organisation of an entire class of substances, as seen in Figure 2, which we refer to as planar-convex arrangement. We think that livers are included in this category. An aircraft P to O is defined as a planar-convex object O in Rn that is defined as a stop up the surface when a set of equivalent hyperplanes P is present, allowing us to acquire an unique closed planar curve in any cross section of an aircraft P to O. This is how we refer to any collection of hyperplanes that are parallel to the O as "convexity planes." We balance two things by coordinating their principle components, and we orient a collection of symmetrical points to the vertices of a square centred on the object's greatest primary constituent by coordinating their principal components. Using this method, we may locate numerous sets of the identical plane, which intersect the object with normals that span throughout a hemisphere equally; we sampled a dodecahedron with 32 vertices to demonstrate the concept. Afterwards, we determine whether main plane (x, y, or z-plane) is more successful in mapping the parallel planes, as indicated by equations (1) and (2);

$$S(p) = \frac{\int_C^1 xV_r(p,x)\mathrm{d}x}{\int_C^1 xV_r(p,x)\mathrm{d}x}. \quad (1)$$

Each plane's link to the liver, as well as the regular number of associated workings, are next thoroughly analysed. This approach will identify the smallest possible sum of average mechanisms in the axis/plane of the two compounds. As a result, the location of matching convexity planes P between two identical objects (as previously specified)—liver segmentation and testing from CT scans of sick patients, and cancer from CT scans of the same patients—can be established. [28] The surface of each convexity plane is sampled using a user-defined number of

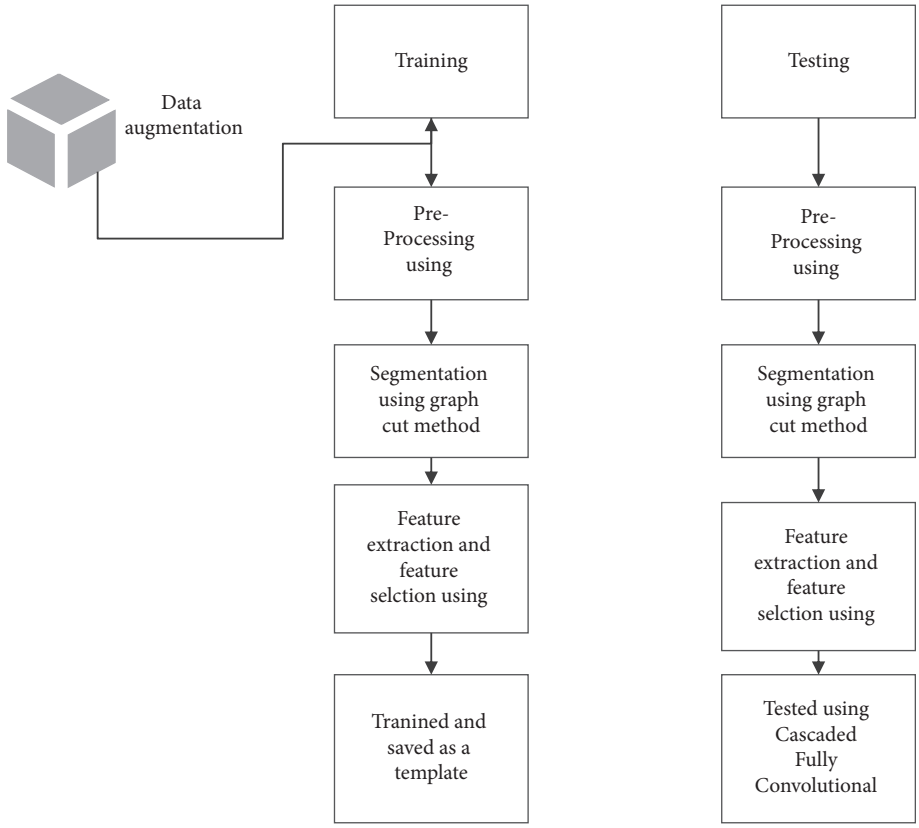


Figure 1: Proposed workflow.

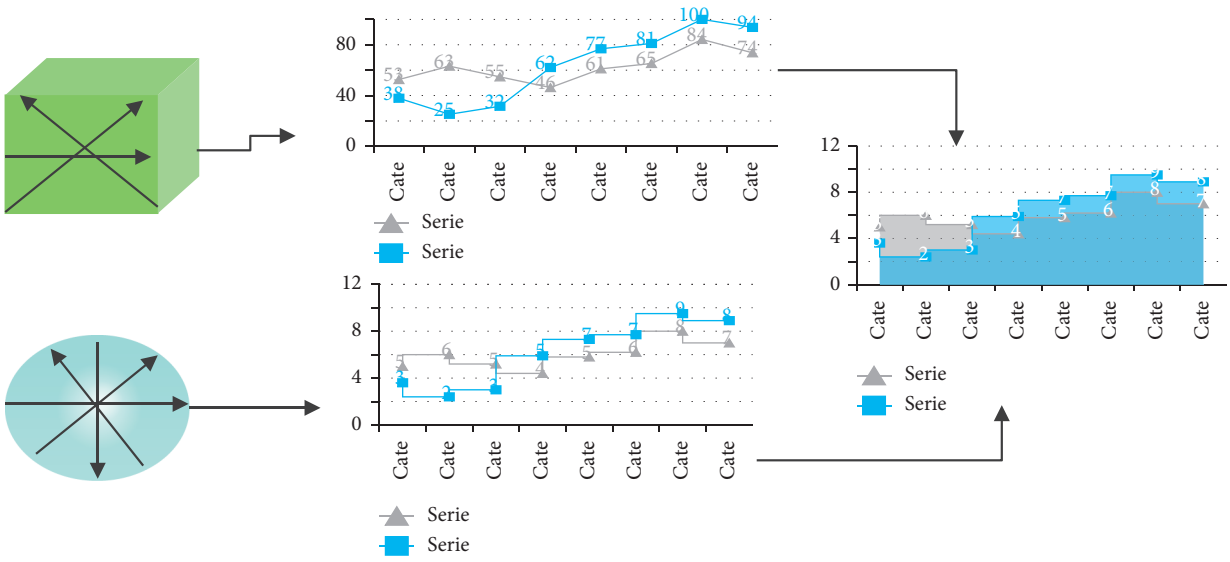


Figure 2: An example of the form descriptor on 2-D closed curves. The seed is a circle that intersects the objects (here a square and a wider circle) (here a square and a larger circle). Provided a suitable parameterization of the two closed curves, their point-to-point local discrepancy can be measured.

divisions, and the points of these partitions are projected onto the surface of the goal, with each partition representing a point on the surface of the target. During the demonstration, you will be shown an example of liver parameters. When comparing two items or lives, these assessments, also known as "parameterization points," are made in relation to one other. In order to study the outcomes, we compute the form function S at each parameterization point and transfer its value onto the surfaces of each entity so that the consequences may be seen. When S is generated from the coordinated preparation surface, it is averaged at both ends, and it is normalised between zero and one on the scale of 0 to

one. Using the most elementary parametrization, this is seen in Figure 3.

3.3. Segmentation of Liver Tumors and Lesions. The first liver segmentation reveals regions on the surface of the liver that have an uncertain structure based on the surface points that were matched to the training results in the first step. S has a cutoff of 0.5, and component analysis is used to give unique marks to each uncertain location, allowing for any degree of intra-patient variability to be accommodated in the study. Because the livers were largely segmented by the original approach, the seeds were placed in the centre of the label, and a rapid marking level was used to "crease" the segmentation based on the sigmoid of the CT images [29], the seeds were placed in the centre of the label. A geodesic contour with dynamic geometry refines the segmentation. The technique is repeated until the volume changes by S 0.5 or until the volume changes between iterations. To characterise timid hepatic masses, a graph-cut method segmenting the liver is utilised, as recommended by the process, to segment the liver. In their simplest version, the graph cuts are affected by the shrinking bias issue, which is especially problematic for the segmentation of enlarged and tiny structures such as blood arteries and some tumour shapes. Tumors and veins are quite diverse from case to case, and the segmentation of abdominal organs with formations has improved as a result of the diagram cuts. Tumors, on the other hand, are often elliptical and curved [30]. It is necessary to compute the tumour vessels and blobs using equations (2) and (3).

$$E_{\text{vessels}} = -\ln \ \ln \max_{\sigma}\left(\sigma^2 v(p,\sigma)\right),$$
$$\text{with } v = \left\{|\lambda_2| + \lambda_1, \text{ if } \lambda_1 < 0\ |\lambda_2| - \frac{\lambda_1}{4}, \right. \tag{2}$$
$$\text{if } \lambda_2 < 0 < \lambda_1 < 4|\lambda_2|.$$

$$E_{\text{blobs}} = -\ln \max_{\sigma}(w),$$
$$\text{with } \lambda_3 > 0; \tag{3}$$
$$w = e^{-(\lambda/\lambda 3^{-1})}.$$

We have reduced the size of the picture by segmenting it using limits for increased vasculature, tumour opacity, and Hessian shape. This allows us to emphasise tiny elongated veins and circular tumours on several scales using our segmented image. The Hessian's values ($p1 > p2 > p3$) at point p highlight major type restrictions that may be used to enhance vascular segmentation and lower the number of false-positive tumours detected. In the graph description, the following force conditions are utilised to describe the graph. Improved hepatic arteries were removed before to tumour segmentation in order to reduce the number of false-positive tumour detections. By standardising the overall volume of the liver, the aggregate quantity of tumours was computed for each patient in order to measure the pressure exerted by tumours and follow the progression of metastatic hepatic cancer [31].

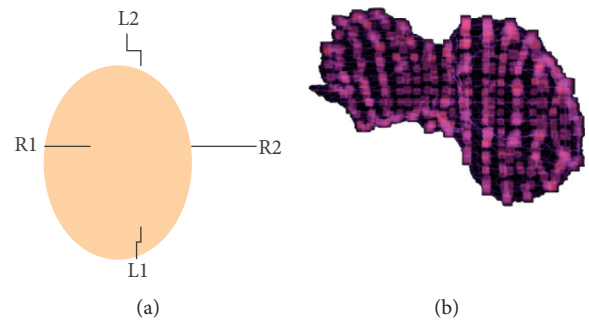


Figure 3: Parameterization points are highlighted as small cubes on the surface of a liver with an irregular shape. These points allow point-to-point correspondence between two shapes.

3.4. Correction of Liver Segmentation. We have reduced the size of the picture by segmenting it using limits for increased vasculature, tumour opacity, and Hessian shape. This allows us to emphasise tiny elongated veins and circular tumours on several scales using our segmented image. The Hessian' values ($p1 > p2 > p3$) at point p highlight major type restrictions that may be used to enhance vascular segmentation and lower the number of false-positive tumours detected. In the graph description, the following force conditions are utilised to describe the graph. Improved hepatic arteries were removed before to tumour segmentation in order to reduce the number of false-positive tumour detections. By standardising the overall volume of the liver, the aggregate quantity of tumours was computed for each patient in order to measure the pressure exerted by tumours and follow the progression of metastatic hepatic cancer [31].

$$E(A) = E_{\text{data}}(A) + E_{\text{enhance}}(A) + E_{\text{shape}}(A) + E_{\text{boundary}}(A), \tag{4}$$

$$E_{\text{data}}(A) = -\sum_{p\in O}^{x} x \ln\left(\frac{\sqrt{P(I_p|O)}}{\sqrt{P(I_p|O)} + \sqrt{P(I_p|B)}}\right) - \sum_{p\in B}^{b} P \ln\left(\frac{\sqrt{P(I_p|B)}}{\sqrt{P(I_p|O)} + \sqrt{P(I_p|B)}}\right),$$
$$E_{\text{boundary}}(A) = \sum_{\{p,q\}\in N_p}^{b} p \exp\left(-\frac{(I_p - I_q)^2}{2\sigma^2}\right)\frac{1}{d(p,q)}. \tag{5}$$

Also included in equation (4) through equation (6) are the voxel intensity and probability of artefacts, as well as the surrounding field, Euclidean distance, and the normal fluctuations in image noise, among others. New language in this formulation refers to the local notion that punishes voxels that do not adhere to the dissimilarity in sharing of better tumours with stable liver parenchyma models, as defined by training. Our index encourages darkish spots within the liver to be identified as tumours since the liver is a better option than cancer in terms of survival. During training on different liver cancers, the relationship between the healthy (background) liver and the diseased (object) liver alters as a result of the training.

$$E_{\text{enhance}}(A) = \sum_{\{p,q\} \in N_p}^{q} x \frac{1}{1 + \left(\left(I'_\delta - I^*_i \right)^2 / 2\sigma^2 \right)^2} \tag{6}$$

where the value of the intensity is specified, and the value of the intensity at the context is specified, the intensity (B). We believe that the surgery is not intended to segment the hepatic vasculature since the improvement is unusual in our circumstances. The traditional geodesic active contour model was utilised to simulate the minor segmentation of tumours in order to maximise their segmentation using a speed spread parameter of five and a curvature parameter of two and a half. By normalising the overall volume of the liver, the total volume of tumours was estimated for each patient in order to measure the pressure exerted by tumours and follow the progression of metastatic hepatic cancer. The absolute difference between the tumour burdens estimated manually and those computed automatically is used to calculate the tumour burden error. The effects of artificial Gaussian noise and body rotation on an axial flat surface were reported and compared with ground reality in order to investigate the reproducibility of estimating tumour burden under the influence of image noise and patient location variations in order to research the reproducibility of estimating tumour burden. When it comes to accentuating circular, multi-scale tumours, the Hessian type is required. Hessian' principles provide certain form limitations that may be used to enhance tumour division while simultaneously reducing the number of false positive tumours. Eqn. 7 has the following energy terms, which are shown graphically in the diagram.

$$E_{\text{shape}} = -\ln \max_{\sigma}(w). \tag{7}$$

We believe that the surgery is not intended to segment the hepatic vasculature since the improvement is unusual in our circumstances. The traditional geodesic active contour was utilised to simulate the minor segmentation of tumours in order to maximise their segmentation with a pace propagation parameter of five and a curve parameter of two and five, respectively. When the entire liver volume was normalised, the total volume of tumours was computed for each patient in order to measure tumour pressure and follow the progression of metastatic hepatic cancer. For the tumour burden error, the absolute difference between the manually computed tumour load and the automatically calculated tumour load is calculated. On the axial level surface, the fake Gaussian cacophony and body rotations were recorded and compared with the ground reality in order to investigate the repeatability of assessing tumour load in the presence of picture noise and patient location alterations.

3.5. Tumor Features and Classification. A set of 157 characteristics is automatically analysed for individual tumor applicants to classify detections. This involves the scale, development, 3D forms, and 3-D texture as seen in Table 1.

The collection of functions in Table 1 was too old to preserve the ideal combination of components to separate accurate positive detections from false-positive detections (TP) because of the large number of classification characteristics that were employed (FP). Due to the fact that various skin textures might overlap and connect together, the classifier must identify the most insightful and distinct characteristics. We can pollute or impair the specific details found in these features if we quantify correlations between exercise samples, which may result in low classification precision. If we quantify correlations between exercise samples, which may result in low classification precision, we can pollute or impair the specific details found in these features. We have conducted tests with a collection of functions, using the methods of least redundancy and maximum application, in this regard (mRMR). mRMR is a feature selection tool that is state-of-the-art in the field of biomedical data processing. Selecting features based on common knowledge and reducing duplication between attributes according to the maximal statistical dependency criterion are two of the benefits of using this method.

3.6. From AlexNet to U-Net. Using a totally convolutional network design for semantic segmentation, Long et al. [32] developed the first such architecture. To create dense predictions by pixels, the researchers use a fully coevolutionary layer structure to replace the last wholly linked layers of a classification network, such as AlexNet, with entirely coevolutionary layers. For the final entirely coevolutionary layers to be adjusted in order to accommodate the input measurements, The AlexFCN (Fully Convolutional Network) improves upon the prior work by allowing full-size medical slices to be projected pixel-wise rather than patch-wise. Using 3D CAD data sets, the AlexFCN training curves (without combining classes) were created. The convergence of all training curves to a stable state occurred quickly when the training and assessment overlapped. AlexFCN has a considerable excess of class equilibrium in both training curves, with Dice overlaps in liver examination exercise knowledge of 90 percent and accidents of 60 percent, respectively, in AlexFCN.

When it comes to examination occasions, the lesion Dice of 24 percent is equivalent to a bad result. It asserted that the class balance was not required in order to resolve their problem with natural picture segmentation. Using AlexNet weights trained on actual photos, for example, might explain why the model was utilised pre-trained in the first place. For training and testing photos, data from ImageNet is utilised.

Table 1: Automated tumor features.

3D feature	Descriptor	Explanation
Tumor volume	Size	Volumetric size
Tumor diameter	Size	Linear size
Tumor size ratio	Shape	
Tumor binay elongation	Shape	Rato of the size of bounding box and real size
Tumor intensity	Shape	Enhancement of tumor region
Edge intensity	Enhancement	Enhancement of healthy region
Cluster	Enhancement	Skewness
Prominence	Texture	Skewness
Edge cluster shade	Texture	Skewness
Correlation	Texture	Complexity
Energy	Texture	Complexity
Entropy	Texture	Roundness
Tumor blobness measure	Texture	Heterogeneity
Inertia	Texture	Heterogeneity
Edge inertia	Texture	Heterogeneity
Tumor inverse difference	Texture	Heterogeneity
Edge inverse difference	Texture	Heterogeneity

Many medical applications, however, need the employment of class balancing because pre-trained networks of real pictures are insufficiently utilised and because the class of attention is less often included in the dataset than the other classes. Preparation and monitoring of Dice for the liver and lesions both improved modestly, with 78 percent of the liver and 38 percent of the lesions being successfully completed on the first attempt. Additionally, the U-Net has a better pattern of skipping connections across different stages in the neuronetwork, in addition to its 19-layer breadth. During the present phase of activations, spatial awareness is accessible in the early stages of the neural network. Spatial information is passed to semantic information at subsequent levels via the neural network, at the price of specific knowledge of the placement of certain structures. Using the original U-Net design, for example, a 388×388 input picture that would otherwise be a bottleneck is reduced to a 28×28 output image. As subsequent stages will merge geographic data from above with neural networks, skip-links will be used to assure later point utilisation and transfer of spatial and semantic data. In later phases, the neural network may make use of semantics and spatial sequencing to make deductions.

3.7. Changes from Fully Convolutional Network to Cascaded Fully Convolutional Network. In the soft mark probability maps P, we have been using the U-Net architecture as a framework. The U-Net design allows for accurate pixel estimation by combining spatial and temporal data into a 19-layer co-evolutionary network architecture—the training U-Net curves in the 3D CAD data set—and merging the results into a single network architecture. In addition, the cumulative lesion segmentation effectiveness has been increased to 53 percent, according to Research Dice. The U-Net has mastered the ability to distinguish between liver and lesion at the same time. One of our most significant innovations is the cascade training of FCN to learn unique features just once during training in order to complete a segmentation assignment, which results in improved segmentation efficiency overall. The approach was developed as a result of the fact that U-Nets and other forms of CNNs recognised the hierarchical structure of the input data. Instead of planning human-crafted face appearances for the separation of distinct tissue kinds, the neural network' stacks of layers are adjusted towards the chosen categorisation in a data-driven manner, rather than by hand. By cascading two U-Nets, U-Net learns from a general CT abdominal scan filtering that is specific to the identification and segmentation of the liver, rather than from a general CT abdominal scan filtering. Figure 2 shows U-Net putting together a filtering process to identify lesions from the liver at the same time as the previous figure. Additionally, the ROI of the liver contributes to the eradication of lesions. We're teaching one network in the abdominal area of the liver, specifically (step 1). It is the only emphasis of this network's research to identify and investigate discriminating traits in liver-background segmentation. After that, we train a second network to segment the lesions in the liver image that we have obtained (step 2). After being segmented in Step 1, the liver is cropped and re-sampled in Step 2 in order to get an input dimension that is suitable for the cascade U-Net. It is possible that the second U-Net will concentrate on learning discriminating properties of the lesion rather than on segmenting the liver history.

Initialize the segmentation process

Begin with features of segmentation image

Let x be feature of the pixels

$y_k = g_m(y_{m-1})$ *be the neuron layers*

While x feature $> y_k$

$y_k = \mathrm{Re}LU(x_m \otimes y_{m-1} + C_m)$

then

$f(y) = m^y(0, y)$

End

where y_m represents a series of convolution operations for each layer. y_k represents the output of layer $m\cdot$ where x_m is

the convolution kernel weight, c_m is the offset value, and $\otimes$ is the convolution operation.

3.8. Effect of Class Balancing. One of the most important steps in FCN training is to balance the needed classes with the class in the data following the pixel frequency of the target. In contrast to [33], we discovered that preparing the system to segment microscopic structures such as lesions is not practicable without class complementary, owing to the substantial class inequity, which is typically in the range of 1% for lesion pixels, and hence not feasible without class complementary. As a result, we have included an additional weighting element in the cross-entropy loss function L of the FCN.

$$L = -\frac{1}{n}\sum_{i=1}^{N} n\omega_i^{\text{class}}\left[\widehat{P}_i \log P_i + \left(1-\widehat{P}_i\right)\log\left(1-P_i\right)\right]. \quad (8)$$

Pi denotes the likelihood of voxel i belonging to the center, $\widehat{P}_i$ represent the position truth. We chose *class i* to be PPi 1i-$\widehat{P}_i\widehat{P}_i$ if $\widehat{P}_i = 1$ (see Figure 4).

4. Experimental Results

We found that the initial segmentation approach was less successful than previously reported [34] because of tumours and other items in our data, as well as the conflicting retrieval of contrast-enhanced pictures. The use of liver-to-liver parameterization in conjunction with active geodesic contour considerably decreased the fraction of volume mistakes in both situations of severe fragmentation failures and those needing modest changes. An example of segmentation from an artifact-free event, a somewhat erroneous segmentation, and a substantial segmentation malfunction are all shown in Figure 5, along with their corresponding type photos and performance during the final repair. With our methods, we were able to enhance the segmentation of crucial instances with tumours while also minimising mistakes in well-secreted livers by a large margin. Since the first and previous segmentation, there has been no arithmetical difference in the segmentation of the liver since the first and prior segmentation.

When manual segmentations were performed on the 14 instances, the usual liver tumour strain was found in 6.6 percent to 9.0 percent and 7.1 percent of the cases when automated segmentations were performed. According to the Wilcoxon rank-sum test, there was no statistically significant difference between the measurements. Figure 6 depicts the change in liver and tumour volume over time, as well as the tumour load, which is significant for many patients.

When we used 3D CRF to our segmentation issue, we were able to demonstrate statistically significant increases in the quality of the segmentation. Because of this, tuning hyperparameters such as 3D CRF requires a significant amount of effort and time. With unintentional search, it is difficult to locate a hyperparameter set that is generalizable to concealed possessions with diverse structure in figure and exterior, such as an HCC lesion. The 3D CRF has also been successfully completed for the treatment of diverse brain

Figure 4: Multiview fusion of proposed cascaded network.

lesions. The introduction of new CRF hyperparameter learning into the training phase was a complete success. When this method is combined with additional words that include prior knowledge of the problem, the CRF's performance for that job may be enhanced.

A Cascaded Fully Convolutional Neural Network for liver tumour detection and segmentation has been proposed for the first time, and it is expected to be widely used in the future. In the system, there is a training phase as well as a testing step for each neural network that is included. The use of data augmentation techniques throughout the training phase helped to increase the overall quality of the CT data that was gathered. It is next necessary to feed the expanded information into the neural network system in order to acquire a qualified framework. This process is known as input data feeding. Our feature extraction strategy comprised the testing of a range of CNN layers in an effort to develop a more effective feature extraction network, which was ultimately successful. This research seeks to overcome the limitations of present spatial 3d information in the identification of neural networks, which are not fully explored at the time of publication. Throughout the Proposal phase, the ideas for the field have been generated from a pyramid structure in order to capture lesions of varied sizes.

The approach is referred to as return on investment (ROI). In contrast, it has been established at this level that a texture classifier can be utilised to distinguish between normal and pathological liver lesions in ROIs collected during the study. Hepatocellular carcinoma (HCC), liver cysts, and hemangiomas irregular hepatic lesions have been distinguished using abstract functions at the classification detection stage, as well as at the classification detection stage and the classification detection stage, respectively. The training phase of this project included a number of iterations that were carried out in order to get a more accurate model structure. During the testing stage, the system was eventually assessed based on the data collected from another batch of CT imaging.

It is obvious from Table 2 that the various segmentation strategies have a variable accuracy rate in terms of classification. Transfer learning using neural network models that have already been trained is a frequent idea in deep learning. When training on a new job, such as medical volume segmentation, neural networks [8] trained on previous tasks, such as a data set for natural image classification, may be used as a starting point for weights of the network to be trained on. The underlying premise of these discoveries is that the initial layers of neural networking for many tasks or datasets uncover a comparable notion to observe crucial systems such as blobs and verges, based on the same theory. When pre-trained models are used, these ideas are not

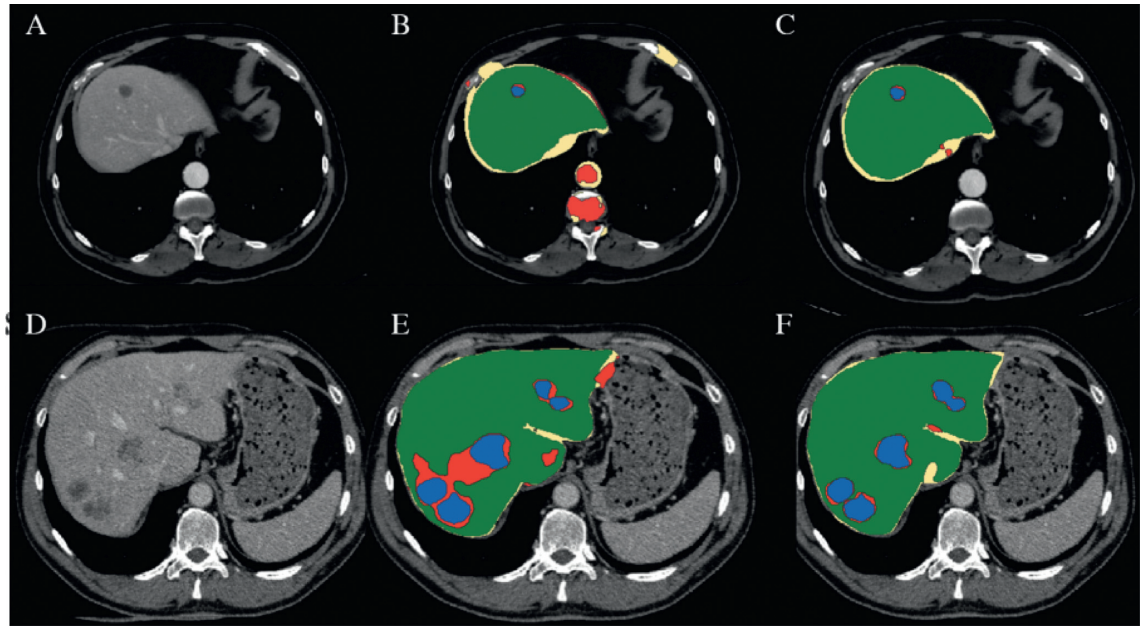


Figure 5: Segmentation of trained and tested features.

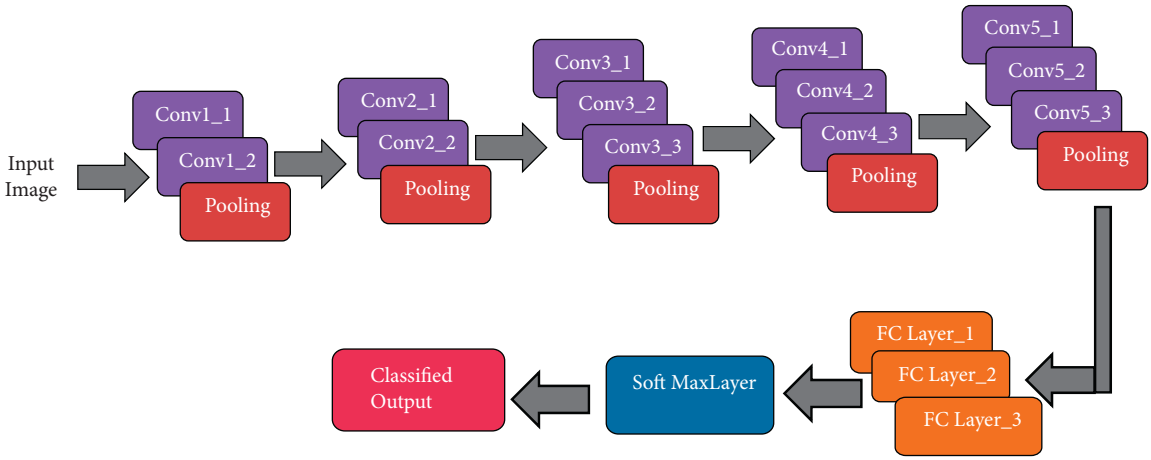


Figure 6: Cascaded Convolutional Neural Network of trained and tested features.

Table 2: Segmented Tumor Parameters using Cascaded Convolutional Neural Network.

Approach	VOE %	RVD %	ASD %	MSD Mm	DICE %
UNET	39.27	87	19.4	119	72.9
Cascaded UNET	12.8	−3.3	2.3	46.7	93.1
Cascaded UNET + 3D	10.7	−1.4	1.5	24	94.3
Proposed	40	89	20	125	89.25

taught from the beginning from scratch. We employ pre-trained U-Net models that have been trained on cell segmentation data to assist our researchers on creating their preparation [7] for our studies, which includes an erudite liver and lesion concept. We have made our taught model on liver and damage segmentation available for download [6].

According to the different current algorithms, as seen in Figure 7, the proposed Unet architecture has the highest rate of accuracy (94.025 percent), as well as the lowest rates of sensitivity and specificity (both 0.5 percent). Due to the longer calculation time required by the other current method, the accuracy rate of the system is diminished.

According to Figure 8, when it came to identifying liver cancer, the sensitivity and specificity were 94.4 and 77.8%, respectively, when compared to other tests. Using an AUC of 0.8070 and a threshold value of 28.35, the sensitivity and specificity for the diagnosis of liver cancer were 83.3 percent

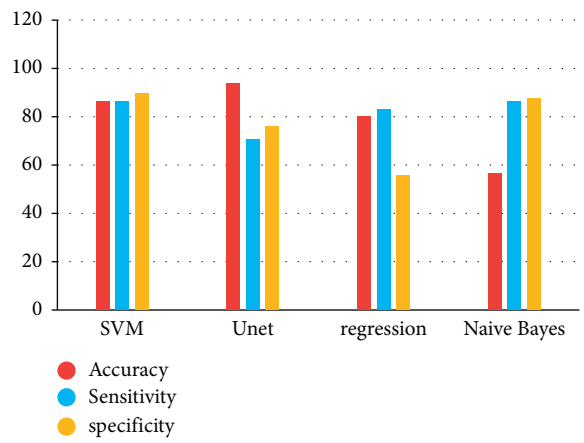


Figure 7: Prediction rate of trained and classified tumor cells.

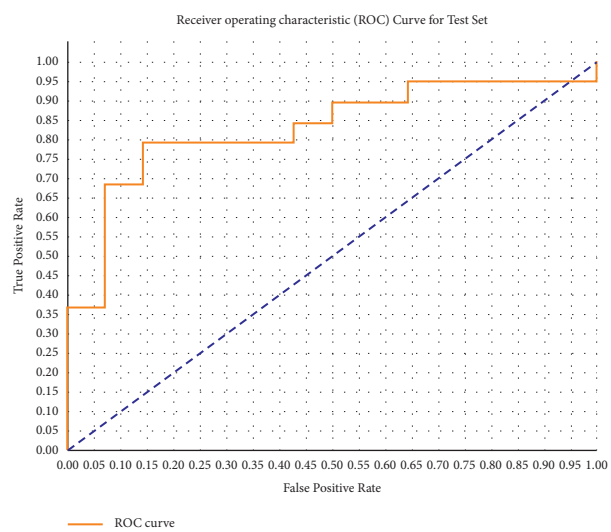


Figure 8: ROC of enhanced UNet architecture with geodesic active contour.

	Alcohol	Cirrhosis	Age	PS	Class	male_0	male_1
Alcohol	1	0.458652	0.162934	0.161536	-0.0403024	-0.442103	0.442103
Cirrhosis	0.458652	1	-0.0014582	0.0224449	0.0375573	-0.253663	0.253663
Age	0.162934	-0.0014582	1	0.152242	-0.146054	-0.172121	0.172121
PS	0.161536	0.0224449	0.152242	1	-0.379708	-0.04661	0.04661
Class	-0.0403024	0.0375573	-0.146054	-0.379708	1	0.0384348	-0.0384348
male_0	-0.442103	-0.253663	-0.172121	-0.04661	0.0384348	1	-1
male_1	0.442103	0.253663	0.172121	0.04661	-0.0384348	-1	1

Figure 9: Classifier stages with respect to routine habitat.

and 77.8%, respectively, showing that the test was both highly sensitive and specific for the disease.

According to the classification stages, the habitats of the people have the most influence on the development of liver cancer. The prediction rate of a person in their everyday life is shown in Figure 9 with regard to their age and environment, respectively. Alcoholism increases a person's risk of developing liver cancer, which is particularly dangerous

for men. This diagnostic tool may be able to detect liver cancer in its early stages, which would aid in the decrease of the fatality rate.

5. Conclusion

Cascade FCNs and 3D CRFs are trained for automated CT volume placement and to co-ordinate the volumetric secretion of illness and injury in order to get optimal results. It is cutting-edge technology that we are proposing to use. We are making our qualifying models available under open-source permission in order to further medical applications in CT data 8. Furthermore, we built and evaluated dense 3D CRF as post-processing measures to inspect a deep medical picture based on learning in order to investigate a deep medical image based on learning. Furthermore, our suggested technique, which involves cascading numerous FCN, has the potential to be applied to many organ segments. The use of additional waterfall FCNs as prospective work ROI for lesions in order to identify the malignancy of the lesions and to do advanced operations. An increase in the use of unfavourable networks may help to enhance the accuracy of segmentation. It is possible to separate the heterogeneous quantities of CT and DW-MRI from different scanners and the specified methods into smaller groups of under 100 apiece. Let us bring this to a close. For automated liver analyses and associated lesions, CFCNs are promising approaches in clinical practise. They are also being studied extensively in medical research [32, 35–42].

Data Availability

The data that support the findings of this study are available on request from the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors extend their appreciation to the Researchers supporting project number (RSP-2021/384) King Saud University, Riyadh, Saudi Arabia.

References

[1] B. Sahin, M. Emirzeoglu, A. Uzun et al., "Unbiased estimation of the liver volume by the Cavalieri principle using magnetic resonance images," *European journal of radiology*, vol. 47, no. 2, pp. 164–170, 2003.

[2] Asia–Pacific Working Party on Prevention of Hepatocellular Carcinoma, "Prevention of hepatocellular carcinoma in the Asia–Pacific region: Consensus statements," *Journal of Gastroenterology and Hepatology*, vol. 25, no. 4, pp. 657–663, 2010.

[3] Y. Yang, C. Li, X. Nie et al., "Metabonomic studies of human hepatocellular carcinoma using high-resolution magic-angle spinning 1H NMR spectroscopy in conjunction with multivariate data analysis," *Journal of Proteome Research*, vol. 6, no. 7, pp. 2605–2614, 2007.

[4] I. Mehmood, N. Ejaz, M. Sajjad, and S. W. Baik, "Prioritization of brain MRI volumes using medical image perception model and tumor region segmentation," *Computers in Biology and Medicine*, vol. 43, no. 10, pp. 1471–1483, 2013.

[5] F. Marcolin and E. Vezzetti, "Novel descriptors for geometrical 3D face analysis," *Multimedia Tools and Applications*, vol. 76, no. 12, pp. 13805–13834, 2017.

[6] A. K. Chaturvedi and P. K. Shukla, "Effective watermarking technique using optimal discrete wavelet transform and sanitization technique," *Multimedia Tools and Applications*, vol. 79, no. 19-20, pp. 13161–13177, 2020.

[7] R. Gupta, P. K. Shukla, and P. Kumar Shukla, "Performance analysis of anti-phishing tools and study of classification data mining algorithms for a novel anti-phishing system," *International Journal of Computer Network and Information Security*, vol. 7, no. 12, pp. 70–77, 2015.

[8] V. Roy, S. Shukla, P. K. Shukla, and P. Rawat, "Gaussian elimination-based novel Canonical correlation analysis method for EEG motion artifact removal," *Journal of Healthcare Engineering*, vol. 2017, p. 11, Article ID 9674712, 2017.

[9] P. K. S. M. Agrawal and A. U. Khan, "Stock price prediction using technical indicators: a predictive model using optimal deep learning," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 8, no. 2, pp. 2297–2305, 2019.

[10] M. G. Linguraru, W. J. Richbourg, J. Jianfei Liu et al., "Tumor burden analysis on computed tomography by automated liver and tumor segmentation," *IEEE Transactions on Medical Imaging*, vol. 31, no. 10, pp. 1965–1976, 2012.

[11] D. Casillas-Perez, D. Pizarro, and D. Fuentes-Jimenez, "The Isowarp: the template-based visual geometry of isometric surfaces," *International Journal of Computer Vision*, vol. 129, pp. 2194–2222, 2021.

[12] B. Preim and M. Meuschke, "A survey of medical animations," *Computers & Graphics*, vol. 90, pp. 145–168, 2020.

[13] B. Seidl, N. Shevchenko, and T. C.. Lueth, "EFFICIENT semi-automatic segmentation of liver-tumors from CT-scans with interactive refinement," *Computer Science*, vol. 216, 2010.

[14] J. Stawiaski, E. Decenciere, and F. Bidault, "Interactive liver tumor segmentation using graph-cuts and watershed," in *Workshop on 3D Segmentation in the Clinic: A Grand challenge II, Liver Tumor Segmentation Challenge*MICCAI, New York, USA, 2008.

[15] W. M. Wells, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz, "Adaptive segmentation of MRI data," *IEEE Transactions on Medical Imaging*, vol. 15, no. 4, pp. 429–442, 1996.

[16] B. P. Brink, R. Veerhuis, E. C. W. Breij, P. van der Valk, C. D. Dijkstra, and L. Bö, "The pathology of multiple sclerosis is location-dependent: no significant complement activation is detected in purely cortical lesions," *Journal of Neuropathology & Experimental Neurology*, vol. 64, no. 2, pp. 147–155, 2005.

[17] M. F. Pasha, K. S. Hong, and M. Rajeswari, "Profiling the features of pre-segmented healthy liver CT scans: towards fast detection of liver lesions in emergency scenarios," in *Proceedings of the 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5169–5173, IEEE, Berlin, Germany, January 2011.

[18] S. Rizzo, F. Botta, S. Raimondi et al., "Radiomics: the facts and the challenges of image analysis," *European radiology experimental*, vol. 2, no. 1, pp. 36–38, 2018.

[19] K. Doman, T. Konishi, and Y. Mekada, "Lesion image synthesis using DCGANs for metastatic liver cancer detection," *Advances in Experimental Medicine and Biology*, vol. 1213, pp. 95–106, 2020.

[20] R. Zhu, C. G. Zhang, Y. Liu et al., "CD147 monoclonal antibody mediated by chitosan nanoparticles loaded with α-hederin enhances antineoplastic activity and cellular uptake in liver cancer cells," *Scientific Reports*, vol. 5, no. 1, pp. 17904–17913, 2015.

[21] S. Ravi and A. K. Singal, "Regorafenib: an evidence-based review of its potential in patients with advanced liver cancer," *Core Evidence*, vol. 9, pp. 81–7, 2014.

[22] W. Abdulrahman, "Diagnosis of liver tumors using image processing," *International Journal of Engineering Research*, vol. 3, no. 4, 2014.

[23] A. Verma and G. Khanna, "A survey on digital image processing techniques for tumor detection," *Indian journal of science and technology*, vol. 9, no. 14, p. 15, 2016.

[24] J. Tang, Q. Sun, J. Liu, and Y. Cao, "An adaptive anisotropic diffusion filter for noise reduction in MR images," in *Proceedings of the 2007 International Conference on Mechatronics and Automation*, pp. 1299–1304, IEEE, Harbin, China, 2007, August.

[25] R. Lin and E. K. Wong, "Morphological operations on images represented by quadtrees," in *Proceedings of the 1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings*, vol. 4, pp. 2203–2206, IEEE, Atlanta, GA, USA, 1996, May.

[26] R. Chandel and G. Gupta, "Image filtering algorithms and techniques: a review," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 10, 2013.

[27] N. Howard, M. Clementino, D. Kim et al., "New developments in mechanisms of prostate cancer progression," in *Seminars in Cancer Biology*, vol. 57, pp. 111–116, Academic Press, Cambridge, MA, USA, 2019.

[28] X. Li, Y. Zheng, W. U. Khan et al., "Physical layer security of Cognitive ambient backscatter Communications for green internet-of-things," *IEEE Transactions on Green Communications and Networking*, vol. 5, no. 3, pp. 1066–1076, 2021.

[29] X. Li, Y. Zheng, M. D. Alshehri et al., "Cognitive AmBC-NOMA IoV-MTS networks with IQI: Reliability and security analysis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 2021, Article ID 3113995, 12 pages, 2021.

[30] G. Khambra and P. Shukla, "Novel machine learning applications on fly ash based concrete: an overview," *Materials Today: Proceedings*, pp. 2214–7853, 2021, ISSN.

[31] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440, Salt Lake City, UT, USA, 2015.

[32] P. k Shukla, J. Kaur Sandhu, A. Ahirwar, D. Ghai, P. Maheshwary, and P. K. Shukla, "Multiobjective genetic algorithm and convolutional neural network based COVID-19 identification in chest X-Ray images," *Mathematical Problems in Engineering*, vol. 2021, p. 9, Article ID 7804540, 2021.

[33] X. Li, M. Zhao, M. Zeng et al., "Hardware impaired ambient backscatter NOMA systems: Reliability and security," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2723–2736, 2021.

[34] J. M. Watt, M. G. Linguraru, and R. M. Summers, "Affine invariant parameterization to assess local shape in abdominal organs," in *Proceedings of the Medical Imaging 2011: Biomedical Applications in Molecular, Structural, and Functional Imaging*, vol. 7965, p. 79650F, International Society for Optics and Photonics, Orlando, Florida, USA, March 2011.

[35] L. Castagneto Gissey, L. Musleh, G. Mariano et al., "Left hepatectomy with concomitant cavo-atrial and biliary tumor thrombectomy for invasive hepatocellular carcinoma: a video report," *Hepatobiliary Surgery and Nutrition*, vol. 7, no. 2, pp. 116–119, 2018.

[36] B. M. Dawant, R. Li, B. Lennon, and S. Li, "Semi-automatic segmentation of the liver and its evaluation on the MICCAI 2007 grand challenge data set," *3D Segmentation in the Clinic: A Grand Challenge*, pp. 215–221, 2007.

[37] T. Heimann, B. Van Ginneken, M. A. Styner et al., "Comparison and evaluation of methods for liver segmentation from CT datasets," *IEEE Transactions on Medical Imaging*, vol. 28, no. 8, pp. 1251–1265, 2009.

[38] M. Erdt and G. Sakas, "March). Computer aided segmentation of kidneys using locally shape constrained deformable models on CT images," in *Medical Imaging 2010: Computer-Aided Diagnosis*vol. 7624, p. 762419, International Society for Optics and Photonics, Bellingham, Washington, 2010.

[39] J. F. Buell, D. Cherqui, D. A. Geller et al., "The international position on laparoscopic liver surgery: the Louisville Statement, 2008," *Annals of surgery*, vol. 250, no. 5, pp. 825–830, 2009.

[40] P. Hu, F. Wu, J. Peng, P. Liang, and D. Kong, "Automatic 3D liver segmentation based on deep learning and globally optimized surface evolution," *Physics in Medicine and Biology*, vol. 61, no. 24, pp. 8676–8698, 2016.

[41] V. Roy, P. K. Shukla, A. K. Gupta, V. Goel, P. K. Shukla, and S. Shukla, "Taxonomy on EEG artifacts removal methods, issues, and healthcare applications," *Journal of Organizational and End User Computing*, vol. 33, no. 1, pp. 19–46, 2021.

[42] N. K. Rathore, N. K. Jain, and P. K. Shukla, "Image forgery detection using singular value decomposition with some attacks," *National Academy Science Letters*, vol. 44, pp. 331–338, 2021.

Hindawi
Journal of Healthcare Engineering
Volume 2022, Article ID 3449433, 15 pages
https://doi.org/10.1155/2022/3449433

Research Article

Improving the Survival Time of Multiagents in Social Dilemmas through Neurotransmitter-Based Deep Q-Learning Model of Emotions

Awais Hassan,[1] Maida Shahid,[1] Faisal Hayat,[2] Jehangir Arshad,[3] Mujtaba Hussain Jaffery,[3] Ateeq Ur Rehman,[4] Kalim Ullah,[5] Seada Hussen,[6] and Habib Hamam[7,8,9]

[1]*Department of Computer Science, University of Engineering and Technology, Lahore 54890, Pakistan*
[2]*Department of Computer Engineering, University of Engineering and Technology, Lahore 54890, Pakistan*
[3]*Department of Electrical & Computer Engineering, COMSATS University Islamabad, Lahore Campus, Lahore 54000, Pakistan*
[4]*Department of Electrical Engineering, Government College University, Lahore 54000, Pakistan*
[5]*Department of Zoology, Kohat University of Science and Technology, Kohat, Khyber Pakhtunkhwa, Pakistan*
[6]*School of Electrical and Computer Engineering, Haramaya Institute of Technology, 138 Diredawa, Ethiopia*
[7]*Faculty of Engineering, Uni de Moncton, Moncton NB E1A3E9, Canada*
[8]*Spectrum of Knowledge Production & Skills Development, Sfax 3027, Tunisia*
[9]*School of Electrical Engineering, Department of Electrical and Electronic Engineering Science, University of Johannesburg, Johannesburg 2006, South Africa*

Correspondence should be addressed to Ateeq Ur Rehman; ateeq.rehman@gcu.edu.pk and Seada Hussen; seada.hussen@aastu.edu.et

Received 8 October 2021; Accepted 28 December 2021; Published 25 January 2022

Academic Editor: Basem M. Elhalawany

In multiagent systems, social dilemmas often arise whenever there is a competition over the limited resources. The major challenge is to establish cooperation among intelligent virtual agents for solving the situations of social dilemmas. In humans, personality and emotions are the primary factors that lead them toward a cooperative environment. To make agents cooperate, they have to become more like humans, that is, believable. Therefore, we hypothesize that emotions according to the personality give birth to believability, and if believability is introduced into agents through emotions, it improves their survival rate in social dilemma situations. The existing researches have introduced different computational models to introduce emotions in virtual agents, but they lack emotions through neurotransmitters. We have proposed a neurotransmitters-based deep Q-learning computational model in multiagents that is a suitable choice for emotion modeling and, hence, believability. The proposed model regulates the agents' emotions by controlling the virtual neurotransmitters (dopamine and oxytocin) according to the agent's personality. The personality of the agent is introduced using OCEAN model. To evaluate the proposed system, we simulated a survival scenario with limited food resources in different experiments. These experiments vary the number of selfish agents (higher neuroticism personality trait) and the selfless agents (higher agreeableness personality trait). Experimental results show that by adding the selfless agents in the scenario, the agents develop cooperation, and their collective survival time increases. Thus, to resolve the social dilemma problems in virtual agents, we can make agents believable through the proposed neurotransmitter-based emotional model. This proposed work may help in developing nonplayer characters (NPCs) in games.

1. Introduction

Artificially intelligent agents are being employed in the field of robotics [1], games [2], entertainment [3], education [4], healthcare [5], customer services [6], and many more. A multiagent system (MAS) is a group of autonomous agents interacting in the same environment to achieve a common goal [7]. In these multiagent systems (MASs), situation of

social dilemmas often arises. As Shaver [8] defined, social dilemmas mean that individuals from a group, society, or culture compete to use limited public goods [9] shared among them. The case of social dilemmas occur in many computational problems such as in competitive structure during file sharing in peer-to-peer systems [10], limited food resources, and their high consumption during simulated survival scenarios [11] and common shared medium among all nodes during bandwidth allocation in telecommunication systems [12].

The capacity to solve social dilemma problems benefits the whole community in the long run. For instance, in Hardin's "Tragedy of the Commons" [13], a social dilemma in the survival scenario, a common pasture, is shared among a community of herdsmen to graze sheep. If each herdsman has a small number of sheep, then the pasture provides plenty of grass to the animals of all herdsmen, which is beneficial for the community in the long run. However, if each herdsman increases his number of sheep for his benefit, the grass is soon scarce in the pasture. The literature suggests that cooperation is necessary among the people to resolve social dilemmas [14–16]. Therefore, to solve the social dilemma among AI-controlled virtual agents, these agents must have believability so that cooperation and coordination are developed among them [17].

The general idea of believability in virtual agents is realistic and human-like characters in virtual worlds. Bogdanovych et al. [18] define a believable virtual agent that is an autonomous agent with its behavior, personality, distinct emotional state, internal goals, and beliefs. This definition explains that personality, emotions, motivation, and social relationships are the key features of believable agents. It suggests that intelligent agents can effectively deal with social dilemma problems when equipped with an empathetic personality through positive emotions, internal motivations, and the capacity to alter their decisions after observing the environment and needs of other agents.

In the literature, believability has been explored initially for virtual agents but only limited to their visual appearance [19]. These works focused on facial expressions [20], motion control [21], hair [22], and dress [23] simulation of virtual agents. Later, many researchers argued that only the physical properties of agents are not sufficient to introduce believability, which can be introduced by making agents rationale that makes goal-oriented decisions [10]. Therefore, the focus was shifted towards the development of models for utility-maximizing rational agents [24].

The rational agents are not adaptable in complex environments as they tend to make self-centered decisions [25]. Hence, these agents must be equipped with emotions, as emotions can affect their goals [26], which in turn alter their actions, thus playing a vital part in decision-making capabilities. The neurological studies suggest that an emotional mind has a substantial contribution to the process of decision-making [27]. Therefore, emotions are a necessity to be included in the rational reactive models for the creation of believable artificial intelligent agents [25]. Literature also suggests that person-specific elements, such as personality [28, 29] and mood [30], also affect the emotion processing mechanism. Therefore, it is also necessary to model the effect of personality on the emotion processing mechanism for the creation of believable virtual agents.

In the last decade, there has been a tremendous advancement in affective computing by introducing various emotional models [31–39] for virtual agents, but these models fail during social dilemmas scenarios. In our opinion, without cooperation between artificial intelligent agents, all the agents' survival as a community is impossible. However, simple rule-based emotions are not enough for multiagents while considering the community's collective survival [38]. Neurotransmitters are the chemicals that control emotions in humans. The development of virtual neurotransmitters in intelligent agents can regulate emotions and help improve the agents' decision-making capabilities. We argue that a neurotransmitter-based emotion modeling in intelligent agents can introduce cooperation and coordination among multiagents and provide collective survival of the community in a virtual world. More specifically, this paper addresses the following research questions.

(i) How do emotions introduce believability in virtual agents?

(ii) How do emotions increase cooperation between agents?

(iii) Does the introduction of emotions by controlling and regulating the virtual neurotransmitters improve the decision-making capability of agents?

The major contribution of this paper is the neurotransmitter-based deep Q-learning model for emotional modeling in virtual agents. According to the OCEAN model, a selfless and selfish personality in virtual agents is established through agreeableness and neuroticism personality traits. Believability based on emotion regulation through dopamine and oxytocin is introduced specifically to the personality of the agent. For the solution of social dilemmas in survival scenario of multiagents, cooperation is established among the agents using the proposed neurotransmitter-based deep Q-learning model.

We have tested the proposed model through simulation performed in a grid world environment developed in the Unity3D platform. Experimentation is conducted by varying the number of selfless and selfish agents, and agents learn to keep their neurotransmitters in the desired range according to their personality. The agents achieve maximum reward by performing the specific actions that best suit their personality. When we increase the number of selfless agents with a high agreeableness personality trait, they start cooperating with others by regulating their positive emotions through neurotransmitters according to their personality, resulting in improved social dilemmas.

The rest of the paper is structured as follows. Section 2 provides a detailed literature survey of the previous work done in this domain. Section 3 presents the proposed solution discussing the model's architecture philosophy and the working of emotional agent in the environment. Section 4 describes the experimentation for the evaluation of the proposed solution. Sections 5 and 6 consist of Results and Discussion. Section 7 concludes the paper.

2. Related Studies

2.1. History of Emotional Models for Believable Agents. The field of emotion-based believable virtual agents flourished after the research works of Bates [40] and his students [41]. They built emotional agents for the Oz project equipped with reactive capabilities and memory. Moreover, they were also introduced with social relationships and emotions models based on Ortony, Clore, and Collins's (OCC) theory of emotions [42]. The research only focused on the agents from an artistic point of view (i.e., appearance-based). The impact of the believable agent's internal and external motivation on its emotions and behavior was not considered. Sloman [43] tried to remove the deficiencies in Bate's Model by introducing a "broad" emotion model. He presented a design-based approach to develop intelligent and motivated agents. While conducting further research to develop a more flexible architecture for autonomous agents, Sloman developed a toolkit named SIM_AGENT [44] for agent development. Many interactive mechanisms of agents, that is, processing different motives, choosing a motive according to the situation, and then acting on it, were included in that toolkit.

For the simulation of dogs and other creatures, an autonomous architecture for agents was proposed by Blumberg [45]. The agents' behavior was created as independent objects, and a specific behavior was achieved by switching between the agent's goals. Hence, the agent would be in a single emotional state at a time, but this architecture failed to address Ekman's complex, compelling, and comprehensive behavioral model [46].

Cathexis is a computational emotional model for producing emotions and their control on agents' behavior proposed by Velàsquez et al. [47]. Although, this model integrated both cognitive and noncognitive promoters of emotions, it did not consider the impact of personality on the emotional changes. However, this model opened new and versatile paths for implementing emotions in agents.

Loyall [48] recognized that personality and emotions both are necessary for the creation of believable agents. He suggested that extensive computations for introducing emotions must also be believable. El-Nasr et al. [49] proposed PETEEI (pet with evolving emotional intelligence). This architecture models the behavior of pet dog through reinforcement learning. This model provided the agent with a feedback mechanism that allowed the agent to adapt its behavior after learning from its experiences.

In the same way, El-Nasr et al. [34] also presented the model FLAME (Fuzzy Logic Adaptive Model of Emotion). This model was also based on the Ortony, Clore, and Collins's (OCC) theory on emotions. An inductive learning system was used to find the hidden patterns in events and connections among objects. Emotions were created based on the appraisal of events according to the fuzzy rules. This model was a significant step forward in determining how emotions change the behavior of agents. Both PETEEI and FLAME were designed for virtual pets; therefore, personality was not included in these models.

Moreover, to improve military simulations, believability in autonomous virtual human agents has also been explored. For this, Silverman et al. [50] compiled human behavior models (HBMs)/performance moderator functions (PMFs), which filtered hundreds of human behavior models helpful in the implementation of behavioral models for believable virtual humans.

Silverman et al. [51] and Silverman, Johns, et al. [52] presented a model that focused on the effects of cognition, stress, perception, and social processes on emotions and virtual decision-making agents. Moreover, they also explored various methods for implementing existing behavioral models in game engines.

Recently, Yang et al. [53] extended the models proposed by Silverman et al. [51] and Silverman, Johns, et al. [52] by introducing the social learning component. This extended model facilitated learning the relationships between agents. Thus, the decision-making capabilities of virtual agents were improved using this information. In the same way, You and Katchabaw [54] presented a model that helped integrate the different psycho-social models.

2.2. Introduction of Cooperation and Emotions through Reinforcement Learning. Diallo et al. [55] proposed that deep reinforcement learning algorithms can be used to cooperate between two agents to achieve a specific task. The fully observable ping-pong scenario tested different deep reinforcement learning algorithms by teaming up the two agents to play against the hard-coded player. Results showed that deep Q-network with double Q-learning gave the best results in the ping-pong scenario by maximizing the total reward of two agents. Cooperation was achieved to some extent, but there was no communication between the agents.

Broekens et al. [56] proposed an emotional model of joy, distress, hope, and fear using reinforcement learning for a single agent in the maze scenario. They mapped the RL primitives to these four emotions. Fear and hope were mapped to the values of the states, whereas joy and distress were mapped to the error function. They hypothesized that for adaptive behavior learning, agents must have complex emotional feedback from the environment. Experiments were done on a maze scenario simulated in JAVA. The results showed that the function of emotion is useful for the adaptive behavior of the agent.

Sequeira et al. [57] proposed an intrinsically motivated reinforcement learning framework to overcome the agent's perceptual limitation by implicitly encoding the information. Reward functions were evolved through the fitness function of genetic programming, and the best-suited functions were adopted to maximize the reward. Many experiments were done on the grid world foraging environment and Pacman scenarios. Results showed that emotional appraisal signals improved the decision-making capability of the agent.

2.3. Introduction of Cooperation for Handling Social Dilemma in Agents. Introducing cooperative behaviors among multiagents has been a topic of interest among researchers for

solving social dilemmas. Researchers have investigated the introduction of emotions among two agents for establishing cooperation in social dilemma scenarios [11, 58, 59].

Yu et al. [58] proposed a double-layered framework with emotional multiagent reinforcement learning that provided agents with emotional and cognition capabilities to induce cooperation. Intrinsic rewards were used to learn the inner layer of the framework, where one emotion emerges as a dominant factor from the emotional processes of the agent. In the outer layer of the framework, the emergent emotion was used as biased reinforcement signals to learn the related cognition and behavioral changes. Experimental results showed that the agent's heterogeneities and different network topologies also had a noteworthy impact on the learning behaviors of the agent.

Huang et al. [59] modeled a dynamic network whose weights evolve with the game's strategy for modeling the effect of cooperation. The hypothesis was that different agents perceive the puzzle of social dilemmas differently. Tests were conducted on two spatial games (prisoner's dilemma and snowdrift game). According to the results, for a small temptation to defect in the game, high network evolution strength was needed for cooperation. The vice versa was also correct.

Leibo et al. [11] introduced sequential social dilemmas to analyze the dynamics of agents' learning policies using deep Q-network. Previously, in Matrix games like the prisoner's dilemma choice to defect or cooperate was treated as a single action. However, in social dilemmas of the real-world, cooperation can be learned after devising the whole policy. For this, the authors tested their hypothesis on two games (fruit gathering and wolf pack hunting game). The experimental results showed that cooperative policies were easier to learn in the fruit-gathering game than wolf pack hunting. As in the wolf pack hunting game, a lot of coordination was needed for learning the cooperative policy. In this research, the personality and emotions of the agents were not taken into account.

A research matrix for comparison of related works is presented in Table 1, where we categorized the existing works on the basis of reinforcement learning, social dilemma in MAS, use of emotions in social dilemmas, and cooperation of multiagents.

3. Proposed Methodology

3.1. Architecture Philosophy. The proposed solution is based on the philosophy shown in Figure 1.

In humans, neurotransmitters are called chemical messengers in the brain. Different levels of these neurochemicals control emotions. However, what emotion has to be expressed depends on the specific personality of humans. Similarly, by varying the levels of virtual neurotransmitters in agents, particular emotions can be acquired according to the agents' personalities. These personality-specific emotions will lead to believability in virtual agents, which is necessary to solve social dilemmas with limited resources.

3.2. Working of Agent in Environment. The model's architecture diagram (Figure 2) is inspired by Barto et al. [29] concept of extrinsic and intrinsic motivation of agents in reinforcement learning.

According to the architecture diagram (Figure 2), the emotional agent can observe the environment. After receiving the knowledge of the environment, the agent performs specific actions in the environment and its state changes. On performing these actions, the agent receives some reward according to extrinsic and intrinsic motivation based on the agent's state. The cumulative reward is calculated by combining rewards based on extrinsic and intrinsic motivation. This cumulated reward is also fed into the agent's brain, along with the state of the environment. The brain processes this information and decides the best action for the agent to take in the environment. This process continues, and the agent tries to maximize its reward by performing the best actions in the environment.

3.3. Emotional Agent

3.3.1. Emotional Agent. The environment is a fully observable grid world consisting of certain virtual agents with different personalities and food reservoirs to gather food. Information about the environment is passed to the brain of the agent in the form of state space. The state-space consists of the location of all food reservoirs, available food in these reservoirs, the total number of alive agents in the environment, the position of these agents, the food level of all these agents, and time passed in the environment. Agent's internal information, that is, its food level and the levels of its neurotransmitters, are also a part of state space.

3.3.2. AGENT Actions. Every agent is allowed certain specific actions that it can perform in the environment. With the actions of up, down, right, and left, it can move in the environment or stay in its position by not act. It can also eat from the food reservoir and share food with the other agents. In short, an agent can perform any one of the total 7 actions available in the action list. The action list is{No action, Up, Down, Left, Right, Eat, and Share}.

3.3.3. Rewards System Based on Extrinsic Motivation. In reinforcement learning, there are two types of rewards.

(i) Rewards based on extrinsic motivation

(ii) Rewards based on intrinsic motivation

Psychologists have distinguished between extrinsic and intrinsic motivation. Extrinsic motivation is defined as propelling us to do a task based on some particular rewarding outcome. Whereas intrinsic motivation drives us to do a job because it is inherently enjoyable. In short, behavior driven by external reward is extrinsic motivation, whereas the behavior driven by internal reward is intrinsic motivation. Extrinsic motivation arises after observing the rewards present in the environment. Intrinsic motivation of the human emerges from the person's personality [60] and

Table 1: Research matrix for comparison of related works.

Research work	Reinforcement learning	Social dilemma in MAS	Emotions in social dilemmas	Cooperation
Bates [40, 41]	✗	✗	✗	✗
Sloman [43, 44]	✗	✗	✗	✗
Blumberg [45], El-Nasr et al. [34, 49]	✗	✗	✗ (Emotional virtual pets)	✗
Silverman et al. [50–52]	✗	✗	✗ (Human behavioral models in game engines)	✗
Huang et.al [59]	✓	✓ (Only for two agents)	✗	✓
Diallo et al. [55]	✓	✗	✗	✓
Leibo et al. [11]	✓	✓ (Only for two agents)	✗	✓
Broekens et al. [56]	✓	✗	✗	✗
Sequeira et al. [57]	✓	✗	✗	✗
Our proposed model	✓	✓	✓	✓

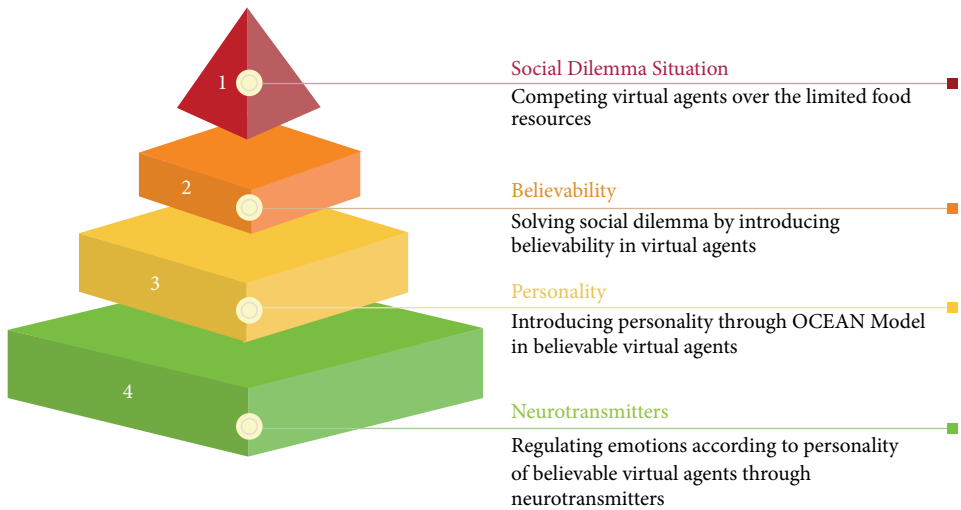


Figure 1: Solving social dilemma situation by introducing believability through personality and neurotransmitters in virtual agents.

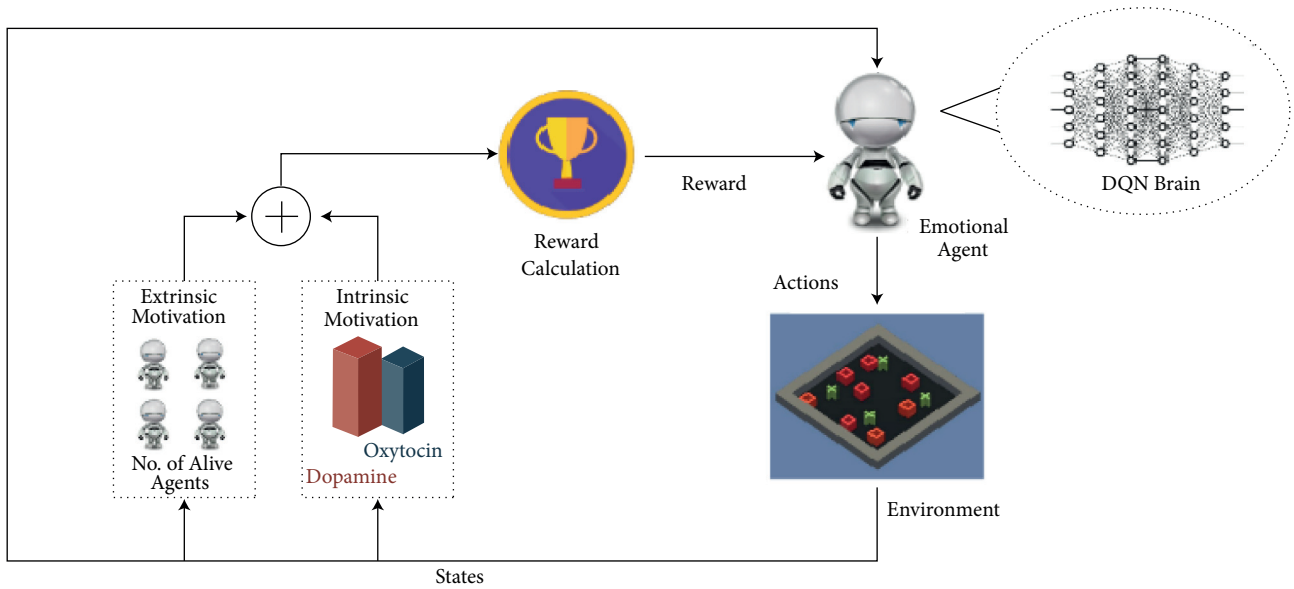


Figure 2: Architecture diagram of the neurotransmitter-based deep Q-learning emotional model with reward calculation based on intrinsic and extrinsic motivation of the emotional agent.

neural networks and dopaminergic systems present in the person's brain [61]. Hence, for the intrinsic motivation in the agent, it is necessary to introduce virtual personality and emotions by controlling the neurotransmitters or brain chemicals in the agent.

3.3.4. Rewards System Based on Intrinsic Motivation. Intrinsic motivations of the agents are inspired by the drive theory of Hull in psychology [62–64]. The specific personality of the agent has some particular emotions. These emotions are achieved by balancing and controlling the neurotransmitters. This balancing of neurotransmitters drives the agent to perform the specific actions related to the agent's personality.

3.3.5. Personality. The personality of the agents is based on the OCEAN model (Figure 3), also known as the Five-Factor Model (FFM) developed by Robert McCrae and Paul Costa [65]. The OCEAN Model consists of big five personality traits: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism.

Openness is an extensive realization of diversity of experience, exceptional ideas, and curiosity. Open to experience, people are willing to seek and strive for new things and are curious intellectually.

Conscientiousness is the likelihood of a person to be dutiful, self-disciplined, and goal-oriented against all measures. It can be associated with how people direct, regulate, and control their stimulus responses.

Extraversion is related to extrovert, action-oriented, and enthusiastic people interacting and engaging with external people and the outside world. The personality trait of agreeableness is marked with kind, considerate, helpful, trustworthy, and generous nature. Agreeable people have a significant concern for social peace, and they have an optimistic perspective of humans. Selfless people fall under this category. Neuroticism is marked with people having negative emotions such as depression, anger, and anxiety. Ordinary situations and minor frustrations are threatening and hopelessly tricky for them. The occurrence of this trait is most likely to make a person selfish.

3.3.6. THE Proposed Personality Model. Studies suggest that people having a high level of Neuroticism and low level of Agreeableness are usually selfish and self-centered [66]. However, people having a high level of Agreeableness and low level of Neuroticism are generally selfless and empathetic [67]. As this paper deals with the social dilemma situation, only those personality traits are required that are most likely to make the agents selfish or selfless. Hence, the personality of Agreeableness and Neuroticism is considered from the OCEAN Model for the emotional agents. These personality trait values remain within the range of [0, 1].

3.3.7. Neurotransmitters. The emotional chemicals involved in the Limbic System [61] are used to determine the agents' emotions. The basic four emotional chemicals (Figure 4) are dopamine, serotonin, oxytocin, and endorphin [68].

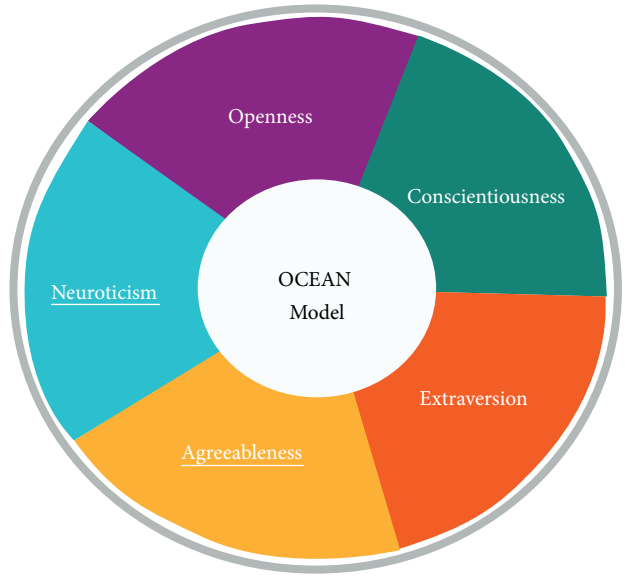


Figure 3: Ocean model of personality.

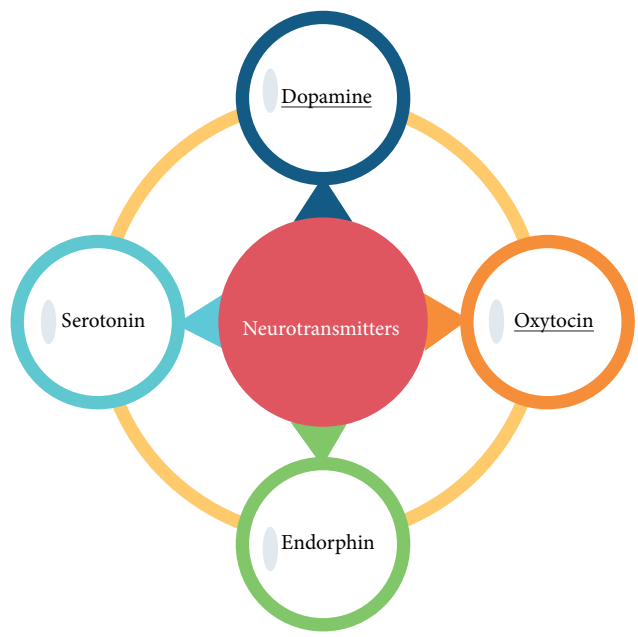


Figure 4: Four types of neurotransmitters.

The neurotransmitter that is related to motivation and reward chemical is dopamine. It is released by a small portion of the brain, hypothalamus, located at the base of the brain. Drive, focus, memory, and attention are associated with this chemical.

Serotonin is the neurotransmitter known as the happiness hormone. Mood upliftment and relaxation are achieved by the correct portion of the serotonin level in the body. Hypothalamus also generates the bonding hormone known as oxytocin. Social behavior and feelings of calmness and contentment are related to this neurotransmitter. In pain, stress, and fear, endorphins are the neurotransmitters released. These allow us to cope with the pain.

In his book [69], Simon Sinek categorized dopamine and endorphin as selfish chemicals, whereas oxytocin and serotonin as selfless chemicals. For the situation of social dilemma in virtual agents, dopamine and oxytocin from each category are considered. Neurotransmitters are maintained within [0,10], with 0 as the lowest level and 10 as the highest level. This paper shows that high level of virtual oxytocin creates the emotion of love, empathy, and selflessness among agents. Low levels of oxytocin generate the emotions of selfishness. Dopamine represents the motivation and goal achievement chemical.

3.3.8. Neurotransmitters to Agents Action's Mapping. As mentioned earlier, the agent can perform any action available in the action list at a time step. While doing these actions, the level of neurotransmitters within the agent changes. For example, if the agent is hungry and takes those movement actions that will reduce its distance from the reservoir, its dopamine level increases. This level of increase depends on the distance between the agent and the food reservoir. Dopamine level increases more speedily if the specific action makes the agent closer to the reservoir. But this level decreases if some action makes the agent move away from the food reservoir. Since during the movement, the agent does not interact and share food with the other agents; therefore, its level of oxytocin decreases. When an agent performs the eat action, the dopamine level increases more swiftly as it has achieved its goal, but the level of oxytocin remains the same at that time. When the agent performs the share action, both the level of dopamine and oxytocin increase. It is because the agent is performing a selfless action.

3.3.9. Extrinsic Motivation. Every agent has a certain food level, and the agent dies if the food level reaches 0. Since the goal of this research is the collective survival of agents; therefore, extrinsic motivation for the agents is the aliveness of every agent. Thus, extrinsic motivation drives the agent to take actions for the survival of all the virtual agents.

3.3.10. Reward Calculation. According to the value of agreeableness and neuroticism, the agent can have different percentages of selflessness and selfishness in its personality as shown in Table 2. In this model, the selfless agent has the agreeableness of 0.8% and 0.2% of neuroticism. Vice versa is correct for the selfish agent.

As already explained, a high level of virtual dopamine and oxytocin creates the emotion of love, empathy, and selflessness among agents. A low level of oxytocin generates the emotions of selfishness. The emotional state of the agents is shown by the following two functions given below:

$$\begin{aligned} F_{\mathrm{SL}}(\mathrm{Dop}, \mathrm{Oxy}) &= \mathrm{sgn}\left(\frac{|((\mathrm{Dop}/4)-1)| + ((\mathrm{Dop}/4)-1)}{2}\right) * \mathrm{sgn}\left(\frac{|((\mathrm{Oxy}/5)-1)| + ((\mathrm{Oxy}/5)-1)}{2}\right), \\ F_{\mathrm{SF}}(\mathrm{Dop}, \mathrm{Oxy}) &= \mathrm{sgn}\left(\frac{|((\mathrm{Dop}/2)-1)| + ((\mathrm{Dop}/2)-1)}{2}\right) * \mathrm{sgn}\left(\frac{|(1-(\mathrm{Oxy}/4))| + (1-(\mathrm{Oxy}/4))}{2}\right), \end{aligned} \tag{1}$$

where subscript SL represents the Selfless and SF represents the Selfish. Sgn is the signum function, and abs is the absolute function. $F_{\mathrm{SL}}(\mathrm{Dop}, \mathrm{Oxy})$ only returns true when the neurotransmitters have reached the specific level, that is, dopamine > 4 and oxytocin > 5. $F_{\mathrm{SF}}(\mathrm{Dop}, \mathrm{Oxy})$ returns true only when dopamine > 2 and oxytocin < 4.

Reward based on the intrinsic motivation R_{Int} is given as follows:

$$\begin{aligned} R_{\mathrm{Int}} = {} & \text{Value of agreeableness} * F_{\mathrm{SL}}(\mathrm{Dop}, \mathrm{Oxy}) \\ & + \text{Value of neuroticism} * F_{\mathrm{SF}}(\mathrm{Dop}, \mathrm{Oxy}). \end{aligned} \tag{2}$$

This reward function gives a greater positive reward of 0.8 to both selfless and selfish agents if they satisfy their nature by adjusting the levels of their neurotransmitters.

Reward based on extrinsic motivation R_{Ext} whenever any agent dies, is given as follows:

$$R_{\mathrm{Ext}} = -(\text{Remaining time of the simulation} * \text{discount factor}). \tag{3}$$

Here discount factor is set to 0.1, which ensures that the reward stays within the range of [−1, 0]. This dynamic reward function makes sure to give a greater negative reward if an agent dies during the episode and less-negative reward if the agent dies near the end of the episode. R_{Ext} is given to every agent if an agent dies to motivate the agents for collective survival. The following equation gives the total accumulative reward.

$$R_{\mathrm{Total}} = R_{\mathrm{Int}} + R_{\mathrm{Ext}}. \tag{4}$$

3.3.11. Emotional Brain. Deep Q-network (DQN), which is a method of deep reinforcement learning (DRL) [70], is used to develop the brain of every emotional agent because it is a continuous problem with no terminal state [71]. By interacting with the environment at discrete time steps ($t = 0, 1, 2 \ldots$), emotional DRL agents learn different policies. Environment state space is denoted by S which consists of all the available information consisting of all the internal and external information of agents and the location of food reservoirs. Action space A consists of the possible actions an agent can perform, that is, $A = \{\text{No action, Up, Down, Left, Right, Eat, Share}\}$. R represents the reward state. At

Table 2: Values of personality traits for both selfless and selfish agents.

Sr. no.	Personality of agent	Agreeableness	Neuroticism
1	Selfless	0.8	0.2
2	Selfish	0.2	0.8

every time step t, every agent observes a state $s_t \in S$ and selects an action $a_t \in A$. In return, the agent gets a reward of $R_{\text{Total}} \in R$ and moves to a new state $s_{t+1} = S$. The agent's goal is to maximize the reward by finding the optimal policy λ: $S \longrightarrow A$, that is, the mapping of observed states and the action taken by the agent in those states.

To address the task as mentioned earlier of DRL, Q-learning is used. The choice of Q-function, the quality of state-action pair (s_t, a_t), is critical to the success of the Q-learning technique. Deep Q-network (DQN) uses deep neural network (DNN) to learn the Q-functions through iterative updates based on the experience. Neural network with one input layer, three hidden layers, and one output layer is used to approximate the action values $Q = (s_t, a_t, \theta)$, where θ represents the learning parameters of the neural network.

The network's input is the agent's state of the environment containing the information about the location of food reservoirs and other agents, food levels of agents, its neurotransmitter levels, and the current simulation time. The output is the approximate Q-value of every possible action that the agent can take as shown in Figure 5. The equation to calculate the Q-values is given as follows:

$$Q(s_t, a_t, \theta) = R_{\text{Total}} + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}, \theta), \quad (5)$$

where R_{Total} is the immediate reward that the agent gets on choosing the best action that gives the maximum Q-value of the next state represented by $\max_{a_{t+1}} Q(s_{t+1}, a_{t+1}, \theta)$. γ is the discount factor. The algorithm of temporal difference (TD) is used [72]. It enables the agent to update its knowledge on every timestep t. The formula is given as follows:

$$\text{TD}(a_t, s_t) = R_{\text{Total}} + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}, \theta) - Q(s_t, a_t, \theta). \quad (6)$$

Substituting equation (6) in equation (5) makes the following equation:

$$Q(s_t, a_t, \theta) = Q(s_t, a_t, \theta) + \alpha \text{TD}(a_t, s_t). \quad (7)$$

DQN is to minimize the mean squared error of the temporal difference, which is shown above. α represents the learning rate.

4. Experimentation

This section explains the experimentation based on the proposed neurotransmitter-based deep Q-learning computational model. The simulation environment consists of four food reservoirs and two types of agents.

(i) Selfless agents

(ii) Selfish agents

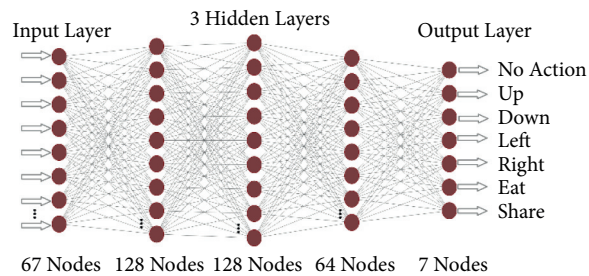


Figure 5: Neural Network used for Q-Learning.

The grid world environment is developed in the Unity3D platform with the gird size of 10×10. Food reservoirs and agents are placed randomly in the environment. Each episode runs for 9 minutes, and after the completion of an episode, the environment resets. The environment also resets when all the agents die before the time of the episode runs out. Each experiment is trained for 100 episodes. DQN brain, implemented in python, is used for the training of the agents. Two experiments are done with altering the number of selfless and selfish agents and checking the effect on the survival time of the agents.

4.1. Food Reservoir. Each food reservoir is initialized with the available food level of 4 (Table 3), which is less than the total food needed by all agents for their survival. It ensures that the situation of social dilemma arises as depicted in Game Theory [73]. Both selfish and selfless agents can consume food from the reservoirs. It is done if the distance between the agent and a particular reservoir is less than 1, and the agent performs the action Eat. Otherwise, the action Eat has no effect.

Whenever the agent takes food from the reservoir, the food level of the agent is increased, and the reservoir storage is decreased. After every minute, the food in every reservoir is regenerated, and the available food is incremented by 0.5. Once the food level is less than 1 in a reservoir, it will not provide food to any agent.

4.2. Selfless and Selfish Agents. Agents are initialized with the food level of five. Following are the seven actions that selfless and selfish agents can perform {No Action, Up, Down, Left, Right, Share, and Eat}. After every minute, the food level of the agents gets decremented by 1. If the food level of any agent is less than 3, it is pushed in an FIFO (First In First Out) queue of needy agents. These are the agents whose food level is less than 3, and they need food from other agents. Both selfless and selfish agents transfer food to the first needy agent in the queue only if its food is greater than 3, as depicted by Maslow's hierarchy of needs [41]. When an agent performs the share action, its food level gets decremented by 0.5, and the food of needy agent (with whom the agent has shared) gets incremented by 0.5. The agent dies if the food level decreases to 0.

Table 3: The initial levels of neurotransmitters and the food level of both types of agents at the beginning of the experiments.

Agents	Food level	Dopamine	Oxytocin	Agreeableness	Neuroticism
Selfless	5	1	4	0.8	0.2
Selfish	5	1	2	0.2	0.8

4.3. DQN Brain. The learning rate of the neural network used for the training of agents is set to 0.0001. Environmental states are passed to the network in batches of 32. The number of nodes in the input layer is 67. First, second, and third hidden layers contain 128, 128, and 64 nodes, respectively. The number of nodes in the output layer is 7. The memory size of each agent to remember the previous states and corresponding actions taken in those states is 100,000. The value of discount factor γ is 0.9.

4.4. Experiment 1. The first experiment (Figure 6) was conducted with eight virtual agents. All were initialized with a selfish personality having agreeableness and neuroticism value 0.2 and 0.8, respectively.

4.5. Experiment 2. The second experiment (Figure 7) was also conducted with eight virtual agents. Out of the eight agents, three agents were initialized with a selfish personality having agreeableness and neuroticism value 0.2 and 0.8, respectively, the same as the first experiment. The remaining five agents were initialized with the selfless personality having agreeableness and neuroticism 0.8 and 0.2, respectively.

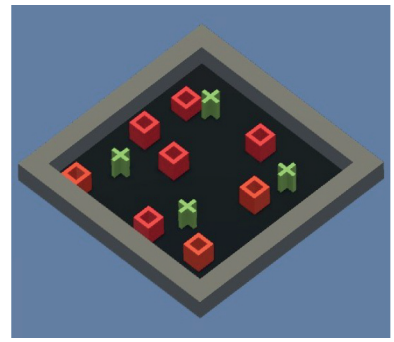


Figure 6: Grid world Unity3D image taken during the training of the first experiment. Plus signs in green show the food reservoirs and red cubes show the selfish agents.

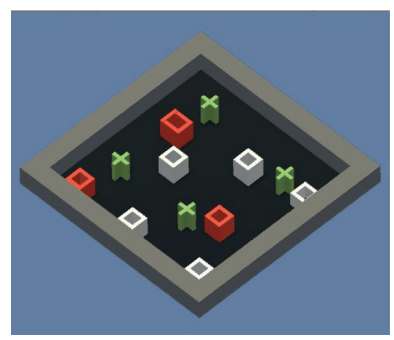


Figure 7: Grid world Unity3D image taken during the training of the second experiment. Plus signs in green show the food reservoirs, red cubes show the selfish agents, and white cubes show the selfless agents.

5. Results

Table 4 shows the survival time of the agents collectively as a community for both the experiments.

Figures 8 and 9 show the Eat and Share actions performed by all the selfish agents during Experiment 1. The x-axis shows the no. of episodes, whereas the y-axis shows the total number of actions performed during a particular episode.

Figures 10 and 11 show the Eat and Share actions performed by all the selfless and selfish agents during Experiment 2. The x-axis shows the no. of episodes, whereas the y-axis shows the total number of actions performed during a particular episode.

Figures 12 and 13 show the Eat and Share actions performed by one selfish and one selfless agent during the 53rd episode of Experiment 2. The x-axis shows the time in seconds on which the specific action was performed. Whereas, y-axis and z-axis show the level of dopamine and oxytocin, respectively, at the particular time the specific action was performed.

6. Discussion

First research question addresses the concept of believability in virtual agents. Believability in virtual agents is based on the personality and emotions of the agents. Furthermore, there are five personality traits, according to the OCEAN Model. From those five traits, Agreeableness and Neuroticism are best suited for the situations of social dilemmas in virtual agents. These two personality traits made the agents selfish and selfless thus contributing to the believability in virtual agents. Moreover, the regulation of emotions according to the personality depends on the neurotransmitters. Dopamine and oxytocin are classified as selfish and selfless neurochemicals, respectively. In this work, we introduce two neurochemicals to introduce emotions and, subsequently, believability.

We performed two experiments to evaluate whether the proposed method introduces believability and how effective it is to solve the social dilemma problems. In the first experiment, all eight agents were selfish, whereas, in the second experiment, five agents were selfless, and three were selfish. Agents were given seven actions {No action, Up, Down, Left, Right, Eat, and Share}. On performing any of those actions, their neurochemicals change. It is evident from Figures 8 and 9 of the first experiment and Figures 10 and 11 of the second experiment that selfish agents performed more eat

Table 4: Collective survival time of the agents for both experiments.

Sr. no.	Time of collective survival (hh:mm:ss)
Experiment 1 (eight selfish agents)	00:06:10
Experiment 2 (five selfless and three selfish agents)	00:07:55

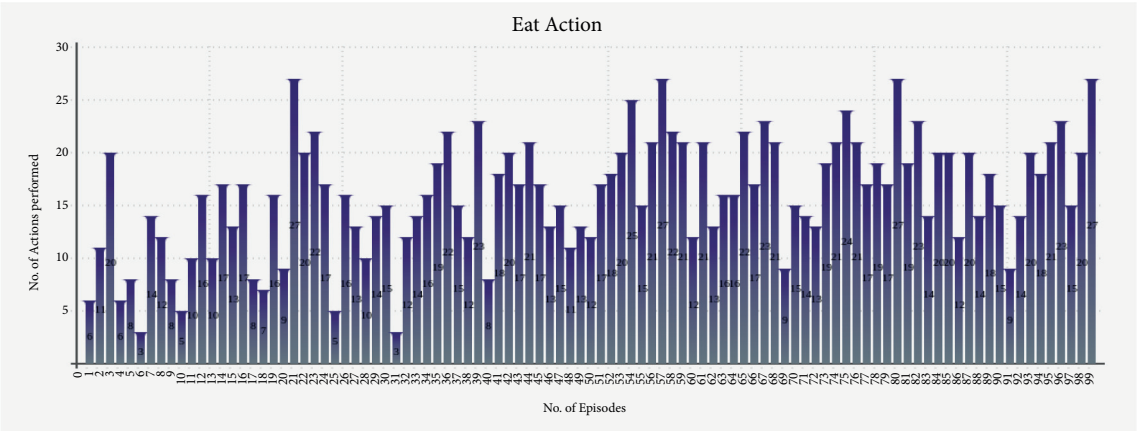


Figure 8: Cumulative Eat actions of eight selfish agents during 100 episodes of first experiment.

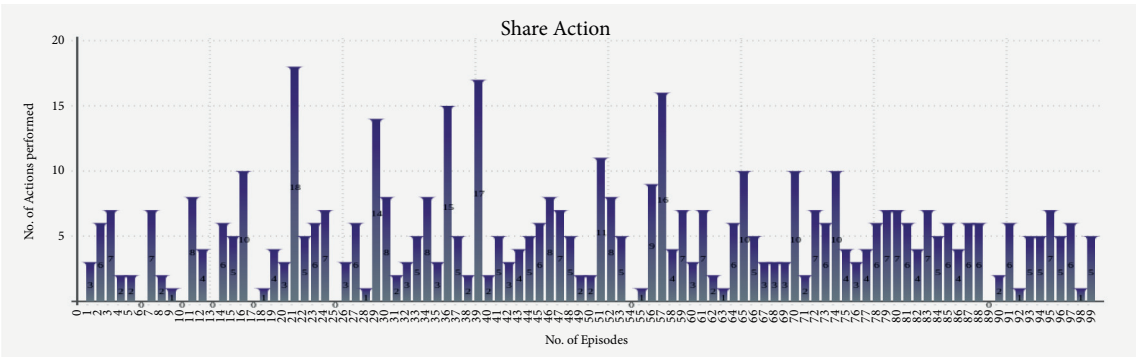


Figure 9: Cumulative Share actions of eight selfish agents during 100 episodes of first experiment.

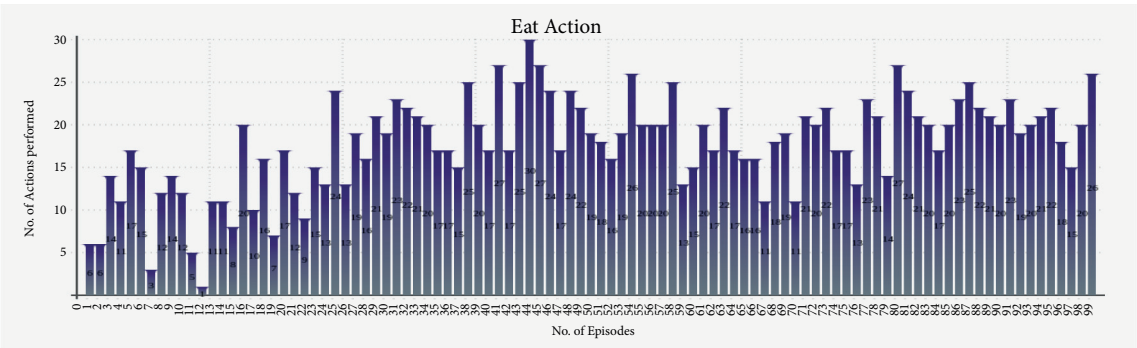


Figure 10: Cumulative Eat actions of five selfless and three selfish agents during 100 episodes of second experiment.

actions and fewer share actions to keep their dopamine level high and oxytocin level low (Figures 12 and 13). However, selfless agents performed more share actions to keep their oxytocin level high (Figure 13). Agents chose those actions that maintained their neurotransmitters in the desired range according to their personality. Therefore, it was concluded that we can introduce believability in virtual agents by regulating the emotions through neurotransmitters according to the agents' personality with the reinforcement learning technique.

Second research question investigates the effect of emotions on cooperation between agents. In the situation of

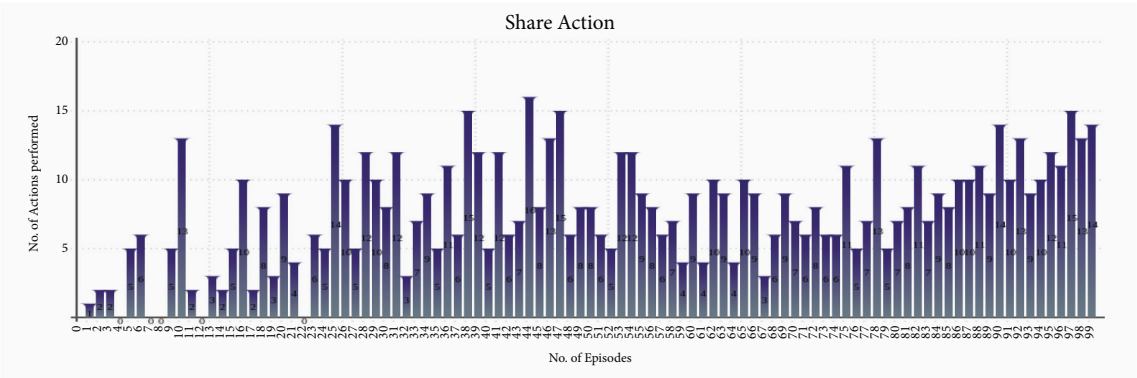


Figure 11: Cumulative Share actions of five selfless and three selfish agents during 100 episodes of second experiment.

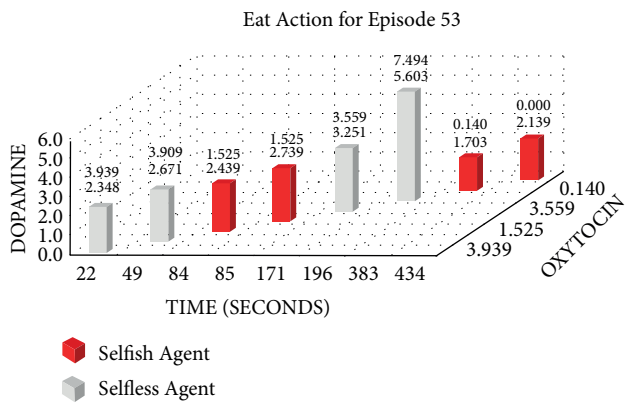


Figure 12: Changes in neurotransmitters of one Selfless and one Selfish agent during the 53rd episode of the second experiment while performing eat action.

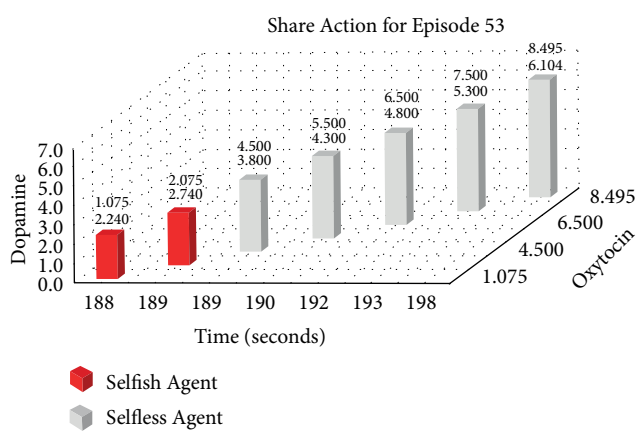


Figure 13: Changes in neurotransmitters of one Selfless and one Selfish agent during the 53rd episode of the second experiment while performing share action.

social dilemmas with limited resources, the situation worsens if people behave selfishly. When people keep their benefits aside and cooperate selflessly, the situation becomes relatively favorable. Therefore, we believed that if we introduce agents with empathetic and selfless personality in the virtual social dilemma situation, the cooperation between the agents can be increased. We also measured the total time for which all the agents were alive (Table 4) while food resources kept the same in both experiments. In Experiment 1, selfish agents ate food for themselves and shared less food with others. Therefore, they could survive collectively for 6 minutes and 10 seconds only (Table 4). However, in Experiment 2, selfless agents ate food from the reservoir and shared food with the needy agents. Therefore, the survival time was increased to 7 minutes and 55 seconds (Table 4). All agents in Experiment 2 survived for 1 minute and 45 seconds more than the eight selfish agents in Experiment 1 with the same food resources.

Hence, it proved that when we increase the number of selfless agents with a high agreeableness personality trait, they started cooperating with others by regulating their positive emotions through neurotransmitters according to their personality. Thus, the situation of social dilemmas improved.

Third research question measures the effect of emotions by controlling virtual neurotransmitters on the decision-making capability of agents. Reinforcement learning is a continuous learning process in which the problem is modeled as Markov decision process (MDP). In our reinforcement learning-based approach, an agent utilizes its previous experience (actions taken in the past) to improve the decisions in the future. The agent with the passage of time discovers which actions give the maximum reward by exploiting and exploring them. Thus, the agent starts to take actions whom Q-values are greater. In our experimentations about reward maximization, selfish agents learnt not to share the food whereas selfless agents learnt to share food with needy agents (who were unable to eat from the reservoir due to limited food) as the training episodes passed. Therefore, through reinforcement learning, agents learnt to keep their neurotransmitters in the desired range according to their personality. Thus, agents were achieving the maximum reward by performing the specific actions that best suited their personality without being explicitly told. This shows that the decision-making capability of agents was improved under reinforcement learning.

In the experiments, we measured the total time for which all the agents were alive (Table 4) while food resources kept the same in both experiments. In Experiment 1, selfish agents ate food for themselves, as shown in Figure 8. The

number of eat actions performed collectively by selfish agents are greater (Figure 8) than the number of share actions (Figure 9). Due to limited food available in the reservoirs, agents were able to survive collectively for 6 minutes and 10 seconds only (Table 4) after 100 episodes of training. After 6 minutes and 10 seconds, first agent died in the environment. However, in Experiment 2, selfless agents ate food from the reservoir as shown in Figure 10. After eating the food, selfless agents also shared the food with the needy agents. It is evident that the number of Share actions in Experiment 2 (Figure 11) is greater than the Share actions performed in Experiment 1 (Figure 9). The survival time was increased to 7 minutes and 55 seconds (Table 4) in this case. All agents in Experiment 2 were able to survive for 1 minute and 45 seconds more than the eight selfish agents in Experiment 1 with the same food resources. It proved that when we increase the number of selfless agents with a high agreeableness personality trait, they started to perform more share actions thus cooperating with others agent. This was because, the selfless agents tried to maximize their reward by regulating their positive emotions through neurotransmitters according to their personality. In return, the situation of social dilemma was improved.

Therefore, we hypothesized that we can introduce emotions by controlling and regulating the virtual neurotransmitters in agents through the reinforcement learning technique to increase the decision-making capability of the agents according to their personality. Two experiments were conducted with varying the number of selfless and selfish agents to evaluate the solution. Both types of agents tried to maximize their reward function by performing actions with the highest Q-value (8). For selfish agents, dopamine level greater than 2 and oxytocin level lower than 4 acquire maximum reward. According to (8), Q-value for Eat action was mostly greater than Share action. Therefore, selfish agents performed more Eat actions for increasing their dopamine level (Figure 12).

Moreover, they avoided sharing food, which resulted in the oxytocin level having a small value (Figure 13). Similar to selfish agents, selfless agents in the second experiment also tried to maximize their reward function. But for selfless agents, maximum reward requires higher dopamine, that is, greater than 4 and higher oxytocin level, that is, greater than 5. Selfless agents tried to keep their dopamine and oxytocin levels high by consuming the food and then sharing it with the needy agents. In each episode, the Q-value of the action Eat was greater for the first 3 minutes; therefore, both selfish and selfless agents consumed food from the reservoir (Figure 12). After 3 minutes, the Q-value of the action Share was greater for the selfless agents only (Figure 13). Agents can share the food only when their food level is greater than 3. To maximize the reward, selfless agents shared food with needy agents (who could not eat from the reservoir due to limited food).

Therefore, agents learned to keep their neurotransmitters in the desired range through reinforcement learning according to their personality. Thus, achieving the maximum reward by performing the specific actions that best suited their personality. Hence, their decision-making capability was improved.

7. Conclusion

To solve the situation of social dilemmas in virtual agents, we proposed a neurotransmitter-based deep Q-learning model for emotional modeling in agents. Agents maintained their neurotransmitter levels by performing specific actions that maximize intrinsic and extrinsic rewards according to their personality. This mapping of actions to neurotransmitters improved the decision-making capability of the agents and developed cooperation between the agents. Experiments showed that selfless agents cooperated with one another, and they survived the social dilemma situation for 1 minute and 45 seconds more than selfish agents. We have concluded that the agents' personality and their emotion regulation through neurotransmitters introduced believability in virtual agents, and selfless agents, in the environment, helped to avoid social dilemma problems that improved the overall survival of the community.

This work opens up a new dimension for emotion modeling in virtual agents. We have chosen a complex social dilemma scenario to see how agents behave in a comparatively large environment. For application point of view, this work can be used for developing nonplayer characters (NPCs) in games. Future directions for extending this research can include extending the experiments using all the four neurotransmitters. Second, we can devise methods that will evolve the personalities of the agents according to the environmental changes.

Data Availability

No data were used to support this study.

Disclosure

Awais Hassan and Jehangir Arshad are co-first authors. The granting agencies did not contribute in the design of the study and collection, analysis, and interpretation of data.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Awais Hassan and Maida Shahid contributed to actualization, validation, methodology, formal analysis, investigation, software, and initial draft. Faisal Hayat, Jehangir Arshad, Mujtaba Hussain Jafferi, Ateeq Ur Rehman, Kalim Ullah, Seada Hussen, and Habib Hamam contributed to actualization, validation, methodology, formal analysis, investigation, and initial draft. All authors read and approved the final version.

Acknowledgments

The authors wish to thank Hina Parveen, Mamona Zahoor, UmeHabiba Saleem, and Zuhha Azhar, the Bs final-year students of Computer Science and Engineering Department, University of Engineering and Technology, Lahore for their

dedication and hard work. The authors also thank Natural Sciences and Engineering Research Council of Canada (NSERC) and New Brunswick Innovation Foundation (NBIF) for the financial support of the global project.

References

[1] Y. Bar-Cohen, "Smart structures and materials 2002: electroactive polymer actuators and devices (EAPAD)," *Electroactive Polymer Actuators and Devices (EAPAD)*, vol. 4695, 2002.

[2] C. Fairclough, M. Fagan, B. Mac Namee, and P. Cunningham, *Research Directions for AI in Computer Games*, Department of Computer Science, Trinity College Dublin, Dublin, Ireland, 2001.

[3] S. Grand and D. Cliff, "Creatures: entertainment software agents with artificial life," *Autonomous Agents and Multi-Agent Systems*, vol. 1, no. 1, pp. 39–57, 1998.

[4] T. Belpaeme, J. Kennedy, A. Ramachandran, B. Scassellati, and F. Tanaka, "Social robots for education: a review," *Science robotics*, vol. 3, no. 21, p. eaat5954, 2018.

[5] D. Isern and A. Moreno, "A systematic literature review of agents applied in healthcare," *Journal of Medical Systems*, vol. 40, no. 2, p. 43, 2016.

[6] T. Araujo, "Living up to the chatbot hype: the influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions," *Computers in Human Behavior*, vol. 85, pp. 183–189, 2018.

[7] M. Wooldridge, *An Introduction to Multi-Agent Systems*, John Wiley & Sons, Hoboken, NJ, USA, 2009.

[8] K. G. Shaver, *Principles of Social Psychology*, Psychology Press, Hove, ES, UK, 2015.

[9] S. Abele, G. Stasser, and C. Chartier, "Conflict and coordination in the provision of public goods: a conceptual analysis of continuous and step-level games," *Personality and Social Psychology Review*, vol. 14, no. 4, pp. 385–401, 2010.

[10] A. B. Tufail, I. Ullah, R. Khan et al., "Recognition of z lotus through aerial imaging and deep transfer learning approach," *Mobile Information Systems*, Article ID 4310321, 2021.

[11] J. Z. Leibo, V. Zambaldi, M. Lanctot, J. Marecki, and T. Graepel, "Multi-agent reinforcement learning in sequential social dilemmas," in *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*, pp. 464–473, São Paulo, Brazil, May 2017.

[12] N. Salazar, J. A. Rodriguez-Aguilar, J. L. Arcos, A. Peleteiro, and J. C. Burguillo-Rial, "Emerging cooperation on complex networks," in *Proceedings of the the 10th International Conference on Autonomous Agents and Multiagent Systems*, vol. 2, pp. 669–676, Taipei, Taiwan, May 2011.

[13] G. Hardin, "The tragedy of the commons," *Science*, vol. 162, no. 3859, pp. 1243–1248, 1968.

[14] K. A. Brekke, K. E. Hauge, J. T. Lind, and K. Nyborg, "Playing with the good guys. A public good game with endogenous group formation," *Journal of Public Economics*, vol. 95, no. 9–10, pp. 1111–1118, 2011.

[15] A. Gunnthorsdottir, D. Houser, and K. McCabe, "Disposition, history and contributions in public goods experiments," *Journal of Economic Behavior & Organization*, vol. 62, no. 2, pp. 304–315, 2007.

[16] S. Gächter and C. Thöni, "Social learning and voluntary cooperation among like-minded people," *Journal of the European Economic Association*, vol. 3, no. 2-3, pp. 303–314, 2005.

[17] S. Pizzutilo, B. De Carolis, and F. de Rosis, "Cooperative interface agents, multiagent systems, artificial societies, and simulated organizations," in *Socially Intelligent Agents*, pp. 61–68, Springer, Berlin, Germany, 2002.

[18] A. Bogdanovych, T. Trescak, and S. Simoff, "What makes virtual agents believable?" *Connection Science*, vol. 28, no. 1, pp. 83–108, 2016.

[19] W. A. Fetter, "A progression of human figures simulated by computer graphics," *IEEE Computer Graph and Applications*, vol. 2, pp. 09–13, 1982.

[20] P. Kalra, N. Magnenat-Thalmann, L. Moccozet, G. Sannier, A. Aubel, and D. Thalmann, "Real-time animation of realistic virtual humans," *IEEE Computer Graphics and Applications*, vol. 18, no. 5, pp. 42–56, 1998.

[21] M. Cavazza, R. Earnshaw, N. Magnenat-Thalmann, and D. Thalmann, "Motion control of virtual humans," *IEEE Computer Graphics and Applications*, vol. 18, no. 5, pp. 24–31, 1998.

[22] A. Daldegan, N. M. Thalmann, T. Kurihara, and D. Thalmann, "An integrated system for modeling, animating and rendering hair," *Computer Graphics Forum*, vol. 12, no. 3, pp. 211–221, 1993.

[23] F. Cordier and N. Magnenat-Thalmann, "Real-time animation of dressed virtual humans," *Computer Graphics Forum*, vol. 21, no. 3, pp. 327–335, 2002.

[24] S. Hassan, N. Tariq, R. Ali Naqvi, A. Ur Rehman, and M. K. A. Kaabar, "Performance evaluation of machine learning-based channel equalization techniques: new trends and challenges," *Journal of Sensors*, vol. 2022, pp. 1–14, 2022.

[25] X. Li, Y. Zheng, W. U. Khan et al., "Physical layer security of cognitive ambient backscatter communications for green internet-of-things," *IEEE Transactions on Green Communications and Networking*, vol. 5, no. 3, pp. 1066–1076, 2021.

[26] M. G. Dyer, "Emotions and their computations: three computer models," *Cognition & Emotion*, vol. 1, no. 3, pp. 323–347, 1987.

[27] D. Goleman, "Emotional intelligence. why it can matter more than IQ," *Learning*, vol. 24, no. 6, pp. 49-50, 1996.

[28] A. Pickering and P. J. Corr, "JA Gray's reinforcement sensitivity theory (RST) of personality," *SAGE Handb. Personal. Theory Assess.*vol. 1, pp. 239–257, 2008.

[29] J. M. Zelenski, R. Baumeister, and G. Loewenstein, *"The Role of Personality in Emotion, Judgment, and Decision Making," Do Emotions Help or Hurt Decision Making? A Hedgefoxian Perspective*, Russell Sage Foundation Press, St., Manhattan, New York, 2008.

[30] R. Neumann, B. Seibt, and F. Strack, "The influence of mood on the intensity of emotional responses: disentangling feeling and knowing," *Cognition & Emotion*, vol. 15, no. 6, pp. 725–747, 2001.

[31] S. K. Haider, A. Jiang, A. Almogren et al., "Energy e flight path model for cluster head selection in next-generation wireless sensor networks," *Sensors*, vol. 21, no. 24, p. 8445, 2021.

[32] J. Dias, S. Mascarenhas, and A. Paiva, "Fatima modular: towards an agent architecture with a generic appraisal framework," in *Emotion Modeling*, pp. 44–56, Springer, Berlin, Germany, 2014.

[33] J. Dias and A. Paiva, "Feeling and reasoning: a computational model for emotional characters," in *Proceedings of the Portuguese Conference on Artificial Intelligence*, pp. 127–140, Covilhã, Portugal, December 2005.

[34] M. S. El-Nasr, J. Yen, and T. R. Ioerger, "Flame-fuzzy logic adaptive model of emotions," *Autonomous Agents and Multi-Agent Systems*, vol. 3, no. 3, pp. 219–257, 2000.

[35] P. Gebhard, "ALMA: a layered model of affect," in *Proceedings of the fourth international joint conference on Autonomous agents and multi-agent systems*, pp. 29–36, The Netherlands, July 2005.

[36] J. Gluz and P. A. Jaques, "A probabilistic formalization of the appraisal for the OCC event-based emotions," *Journal of Artificial Intelligence Research*, vol. 58, pp. 627–664, 2017.

[37] J. Gratch and S. Marsella, "A domain-independent framework for modeling emotion," *Cognitive Systems Research*, vol. 5, no. 4, pp. 269–306, 2004.

[38] S. Jain and K. Asawa, "EMIA: emotion model for intelligent agent," *Journal of Intelligent Systems*, vol. 24, no. 4, pp. 449–465, 2015.

[39] S. Kshirsagar, "A multilayer personality model," in *Proceedings of the 2nd international symposium on Smart graphics*, pp. 107–115, Hawthorne New York USA, June 2002.

[40] J. Bates, "The role of emotion in believable agents," *Communications of the ACM*, vol. 37, no. 7, pp. 122–125, 1994.

[41] A. B. Loyall and J. Bates, "Real-time control of animated broad agents," in *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society*, pp. 664–669, Boulder, Colorado, June 1993.

[42] A. Ortony, G. L. Clore, and A. Collins, *The Cognitive Structure of Emotions*, Cambridge Uni, Cambridge, 1988.

[43] J. Arshad, A. Khan, M. Aftab et al., "Deep deterministic policy gradient to regulate feedback control systems using reinforcement learning," *Computers, Materials & Continua*, vol. 71, no. 1, pp. 1153–11169, 2022.

[44] A. Sloman and R. Poli, "SIM_AGENT: a toolkit for exploring agent designs," in *Proceedings of the International Workshop on Agent Theories, Architectures, and Languages*, pp. 392–407, Montréal Canada, August 1995.

[45] B. M. Blumberg, "Old tricks, new dogs: ethology and interactive creatures," *Massachusetts Institute of Technology*, Cambridge, MA, USA, 1997.

[46] P. E. Ekman and R. J. Davidson, *The Nature of Emotion: Fundamental Questions*, Oxford University Press, Oxford, UK, 1994.

[47] J. D. Velásquez and P. Maes, "Cathexis: a computational model of emotions," in *Proceedings of the first international conference on Autonomous agents*, pp. 518-519, Marina del Rey, CA, USA, February 1997.

[48] W. U. Khan, N. Imtiaz, and I. Ullah, "Joint optimization of NOMA -enabled backscatter communications for beyond 5G IoT networks," *Internet Technology Letters*, vol. 4, no. 2, p. 2020, 2021.

[49] M. S. El-Nasr, T. R. Ioerger, and J. Yen, "PETEEI: a PET with evolving emotional intelligence," in *Proceedings of the third annual conference on Autonomous Agents 1999*, pp. 9–15, Seattle, WA, USA, 1999.

[50] B. G. Silverman, R. Might, R. Dubois, H. Shin, M. Johns, and R. Weaver, *Toward a Human Behavior Models Anthology for Synthetic Agent Development*, https://repository.upenn.edu/hms/84, 2001.

[51] B. G. Silverman, M. Johns, J. Cornwell, and K. O'Brien, "Human behavior models for agents in simulators and games: part I: enabling science with PMFserv," *Presence: Teleoperators and Virtual Environments*, vol. 15, no. 2, pp. 139–162, 2006.

[52] B. G. Silverman, G. Bharathy, K. O'Brien, and J. Cornwell, "Human behavior models for agents in simulators and games: part II: gamebot engineering with PMFserv," *Presence: Teleoperators and Virtual Environments*, vol. 15, no. 2, pp. 163–185, 2006.

[53] H. Yang, Z. Pan, M. Zhang, and C. Ju, "Modeling emotional action for social characters," *The Knowledge Engineering Review*, vol. 23, no. 4, pp. 321–337, 2008.

[54] J. You and M. Katchabaw, "A flexible multi-model approach to psychosocial integration in non-player characters in modern video games," in *Proceedings of the International Academic Conference on the Future of Game Design and Technology*, pp. 17–24, Vancouver, Canada, May 2010.

[55] E. A. O. Diallo and T. Sugawara, "Learning strategic group formation for coordinated behavior in adversarial multi-agent with double DQN," *Lecture Notes in Computer Science*, in *Proceedings of the International Conference on Principles and Practice of Multi-Agent Systems*, pp. 458–466, Tokyo, Japan, October 2018.

[56] J. Broekens, E. Jacobs, and C. M. Jonker, "A reinforcement learning model of joy, distress, hope and fear," *Connection Science*, vol. 27, no. 3, pp. 215–233, 2015.

[57] P. Sequeira, F. S. Melo, and A. Paiva, "Emergence of emotional appraisal signals in reinforcement learning agents," *Autonomous Agents and Multi-Agent Systems*, vol. 29, no. 4, pp. 537–568, 2015.

[58] C. Yu, M. Zhang, F. Ren, and G. Tan, "Emotional m reinforcement learning in spatial social dilemmas," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 12, pp. 3083–3096, 2015.

[59] K. Huang, Y. Liu, Y. Zhang, C. Yang, and Z. Wang, "Understanding cooperative behavior of agents with heterogeneous perceptions in dynamic networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 509, pp. 234–240, 2018.

[60] S. Watanabe and Y. Kanazawa, "A test of a personality-based view of intrinsic motivation," *Japanese Journal of Administrative Science*, vol. 22, no. 2, pp. 117–130, 2009.

[61] S. I. Di Domenico and R. M. Ryan, "The emerging neuroscience of intrinsic motivation: a new Frontier in self-determination research," *Frontiers in Human Neuroscience*, vol. 11, p. 145, 2017.

[62] C. L. Hull, *Principles of Behavior*, Vol. 422, Appleton-Century-Crofts, , New York, 1943.

[63] W. U. Khan, F. Jameel, N. Kumar, R. Jantti, and M. Guizani, "Backscatter-Enabled efficient V2X communication with non-orthogonal multiple access," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 2, pp. 1724–1735, 2021, Feb. 2021.

[64] A. U. Rehman, R. A. Naqvi, A. Rehman, A. Paul, M. T. Sadiq, and D. Hussain, "A trustworthy SIoT aware mechanism as an enabler for citizen services in smart cities," *Electronics*, vol. 9, no. 6, p. 918, 2020.

[65] R. R. McCrae and O. P. John, "An introduction to the five-factor model and its applications," *Journal of Personality*, vol. 60, no. 2, pp. 175–215, 1992.

[66] N. Sohail, J. Arshad, M. Zeeshan et al., "An empirical study on diabetes depression over distress evaluation using diagnosis statistical manual and chi-square method," *International Journal of Environmental Research and Public Health (IJERPH)*, vol. 18, no. 7, p. 3755, 2021.

[67] B. W. Haas, A. Ishak, L. Denison, I. Anderson, and M. M. Filkowski, "Agreeableness and brain activity during emotion attribution decisions," *Journal of Research in Personality*, vol. 57, pp. 26–31, 2015.

[68] T. Nguyen, "Hacking into your happy chemicals: dopamine, serotonin, endorphins and oxytocin," *Huffingt. Post*, vol. 156, 2014.

[69] S. Sinek, *Leaders Eat Last: Why Some Teams Pull Together and Others Don't*, Penguin, London, UK, 2014.
[70] V. François-Lavet, P. Henderson, R. Islam, M. G. Bellemare, and J. Pineau, "An introduction to deep reinforcement learning," *Found. Trends®in Mach. Learn.*vol. 11, no. 3-4, pp. 219–354, 2018.
[71] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, The MIT Press, Cambridge, MA, USA, 2011.
[72] R. S. Sutton, "Learning to predict by the methods of temporal differences," *Machine Learning*, vol. 3, no. 1, pp. 9–44, 1988.
[73] R. Myerson, *Game Theory: Analysis of Conflict*, Harvard Univ Press, Cambridge, MA, USA, 1991.

Hindawi
Journal of Healthcare Engineering
Volume 2022, Article ID 2793850, 13 pages
https://doi.org/10.1155/2022/2793850

Research Article

Zika Virus Prediction Using AI-Driven Technology and Hybrid Optimization Algorithm in Healthcare

Pankaj Dadheech,[1] Abolfazl Mehbodniya,[2] Shivam Tiwari,[3] Sarvesh Kumar,[4] Pooja Singh,[5] Sweta Gupta,[6] and Henry kwame Atiglah[7]

[1]*Department of Computer Science and Engineering, Swami Keshvanand Institute of Technology, Management & Gramothan (SKIT), Jagatpura, Jaipur, Rajasthan-302017, India*
[2]*Department of Electronics and Communication Engineering, Kuwait College of Science and Technology (KCST), Kuwait*
[3]*Department of Computer Science and Engineering, Integral University, Lucknow, India*
[4]*Department of Computer Science and Engineering, Babu Banarasi Das University, Lucknow, India*
[5]*Department of Computer Science and Engineering, Amity School of Engineering and Technology, Noida, Sector-125, Uttar Pradesh, India*
[6]*Department of CSE, School of Engineering and Technology, JLU, Bhopal, M.P, India*
[7]*Department of Electrical and Electronics Engineering, Tamale Technical University, Tamale, Ghana*

Correspondence should be addressed to Henry kwame Atiglah; hkatiglah@tatu.edu.gh

Received 24 November 2021; Accepted 21 December 2021; Published 12 January 2022

Academic Editor: Xingwang Li

The Zika virus presents an extraordinary public health hazard after spreading from Brazil to the Americas. In the absence of credible forecasts of the outbreak's geographic scope and infection frequency, international public health agencies were unable to plan and allocate surveillance resources efficiently. An RNA test will be done on the subjects if they are found to be infected with Zika virus. By training the specified characteristics, the suggested Hybrid Optimization Algorithm such as multilayer perceptron with probabilistic optimization strategy gives forth a greater accuracy rate. The MATLAB program incorporates numerous machine learning algorithms and artificial intelligence methodologies. It reduces forecast time while retaining excellent accuracy. The projected classes are encrypted and sent to patients. The Advanced Encryption Standard (AES) and TRIPLE Data Encryption Standard (TEDS) are combined to make this possible (DES). The experimental outcomes improve the accuracy of patient results communication. Cryptosystem processing acquires minimal timing of 0.15 s with 91.25 percent accuracy.

1. Introduction

An infection's prelude and rapid spread to other regions of the globe attract the attention of the international community. Climate change is also a significant contributor to the rapid spread of illness. The species *Aedes aegypti*, which is widespread in urban areas, is the primary source of the illness, according to the CDC. Zika virus is a mosquito-borne illness that is comparable to dengue fever, west Nile virus, Chikungunya virus, and yellow fever, all of which are spread by mosquitoes. It is believed that mosquito bites are causing the spread of these illnesses and that *Aedes* mosquitoes are the primary vectors of these infections. Humans have had a significant challenge as a result of those, and this has mostly resulted in the transience of many tropical and subtropical nations [1]. It is an intrauterine illness caused by the Zika virus; the symptoms of Zika virus are often moderate fever, joint discomfort, and rashes, which are similar to those of dengue and Chikungunya virus. If an infected mosquito bites a pregnant woman and she becomes sick, the virus may spread across the placenta and affect the fetus. It is believed that pregnant women who are infected with the Zika virus will have neurological problems such as microcephaly and may give birth prematurely to their children. Even males who are infected with the 20 viruses may spread the virus to their sexual partners via anal, oral, or

vaginal intercourse. The increasing population in urban areas has also increased the demand for water portability, which has resulted in people storing water in their homes, causing the *Aedes aegypti* mosquitoes to breed quickly in that climatic condition [2]. The increasing population in urban areas has also increased the demand for water portability, which has resulted in people storing water in their homes. It is believed that mosquitoes that have formed in water that we use in the home are the primary cause of the ailments. In India, the climatic conditions are ideal for mosquito breeding and development.

Preventive methods for limiting the breading of mosquitoes in the infected region are part of the overall awareness campaign in the affected area: keeping the home clean, making sure there is no standing water in or around the house, and making sure any storage water is properly sealed with a lid if there is any. If any of them are discovered to be infected and exhibiting any of the symptoms, they must notify the appropriate healthcare facilities promptly. The individuals in question will be supplied with the medicine they need.

Cryptography is a critical milestone in the development of network security. The term cryptography refers to anything that is concealed or secret. Cryptography is the aim of secret writing with the goal of safeguarding information. The science of breaching cryptosystems, also known as cryptanalysis and cryptology, is the technique of breaking them down by trial and error. Cryptology plays a significant role in the protection of data in computer networks. Writing and solving data using codes are what cryptology is all about, and it includes both cryptography and cryptanalysis.

Cryptography may be divided into three types: asymmetric, symmetric, and hashing. The following sections provide explanations of cryptography [3]. Asymmetric key cryptography is a kind of cryptography in which the key is not shared between two parties. Asymmetric cryptography, often known as public key cryptography, is distinguished by the fact that it uses both a public and a private key. While encrypting the data, the sender sends both the public key and the secret key used to decrypt the data. Only the person who has the secret key may decode the data throughout the decryption process. As a result, it is extremely safe when compared to the symmetric type, yet the time is extremely sluggish. Figure 1 demonstrates the encryption process.

Having control over the data is regarded to constitute secure at the location where it is kept. If a user wants to make use of the advantages of cloud computing, he or she must first choose an appropriate network and then make use of the dispersed resources and arrangements made possible by cloud computing. When it comes to data transfer, the protection of the information is critical [4]. As a result, data that is really necessary should be secured, and obtaining privacy access rights is a major difficulty. Figure 2 represents the data security issue.

1.1. Secureness in Cloud Database. Generally speaking, confidentiality refers to data or information that is not made available to anybody other than its owner, whether that is a person, a device, or a method. CSP is aware of the location or position of data stored by the user and knows where it is stored. There will be certain data that will provide some access assurance to a restricted number of people where the permission has been applied to particular data to secure data confidentiality in this case. Because of the sensitivity of the data that has been gathered in the cloud, some illegal access might result in risks. Assuredness should be provided to customers in the form of a privacy policy to ensure that data is handled appropriately and that mechanisms are in place to guarantee data security in the cloud. It is the responsibility of CSP to implement a variety of measures to ensure data integrity. The CSP informs the client about the kind of data that is being stored in the cloud [5]. As a result, it is necessary for CSP to maintain records about the data, such as the type of data, whether public or private, when and where it is required, the type of Virtual Memory and accumulated, the period of time when it is accessed, and so on, in order to protect the data from unauthorized access and to maintain data confidentiality.

1.2. Data Allocation in Server. We can access our data from any place and at any time thanks to the cloud, which is a superior service provider in our opinion. A risk is related with the location of data collection, which is associated with a higher degree of risk [6] as compared to other sites. The user should be aware of the location of sensitive data storage while an organization is using it, and he or she should have the ability to seek information about the location. At order to avoid confusion, the CSP and client should be in a certain place where the server location and data storage location are both known to the one in control of the scenario. When moving data from one location to another, such as via emails or uploaded photographs to Facebook, there are a number of things to take into account.

1.3. Recovery of Input Data. While transferring data to the cloud, the CSP offers flexibility and ensures that the storage system has accurate information about the data sent. At the very least, a RAID configuration should be maintained in the storage system. The majority of CSPs will have several copies spread among a large number of independent or free servers. The cloud service provider provides the back-end application service, and in the event that a problem happens inside the company, the cloud service provider retrieves the data.

1.4. Secureness in Data Using Encryption Standards. According to a study report from 2100 Indian Business Technology Professionals, data auditing and confidentiality are the most significant obstacles in organizations using cloud technology [7]. The results of a survey conducted by Salt March Intelligence reveal the amount of sensitivity of business professionals to various technologies, including the difficulties they have in embracing cloud apps, infrastructures, customers, and storage. In today's fast-paced company climate, agility, cost savings, and flexibility are all required. The cloud environment provides all of these benefits. Zika virus illness prediction and data security are critical for both

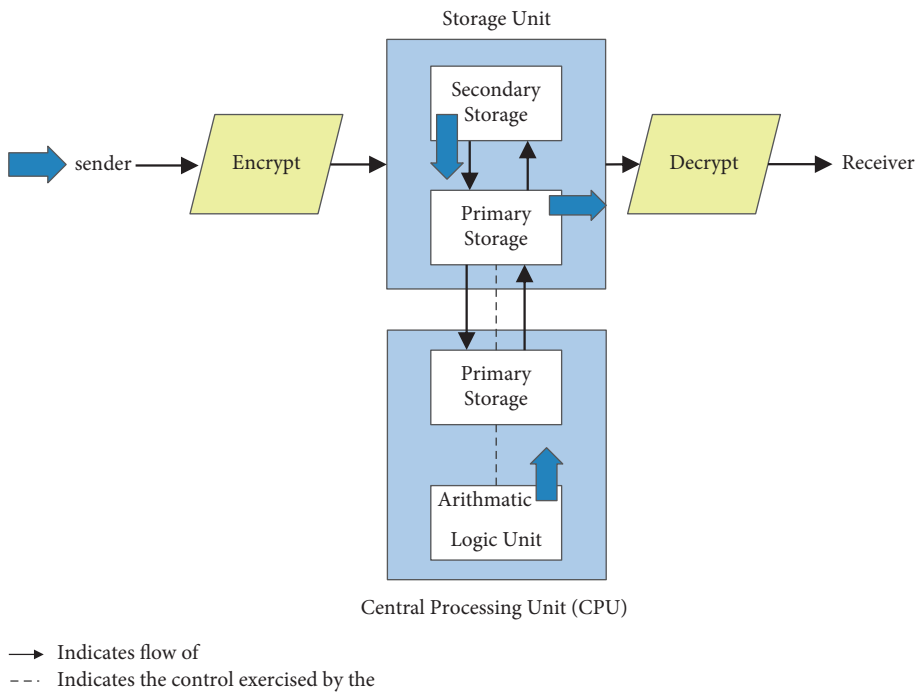


Figure 1: Asymmetric key encryption.

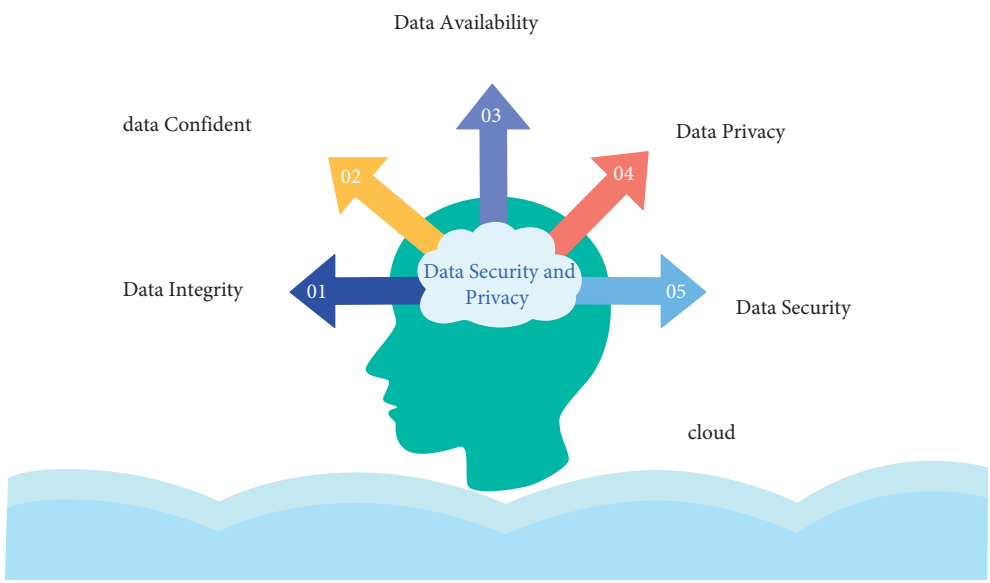


Figure 2: Data security issues.

protecting patients from sickness and safeguarding their personal information. Nowadays, everything is being moved to the cloud, and healthcare is one of the industries that is being moved to the cloud. As a result, more precision in prediction is required; yet, only a limited amount of study has been done in order to anticipate the ZIKA virus. The amount of time it takes to perform a security request may vary depending on the configuration of the system or program being utilized. It is necessary to design a gap prediction and security algorithm in order to triumph over the research and to meet the requirements. In order to improve performance in forecasting the Zika virus with high levels of accuracy, the primary goal of the study activity is to develop new methods of prediction. Because there is no more data available, synthetic data is being generated for the prediction of the Zika virus. The data is divided into two groups based on the results of the machine learning classifier: infected and uninfected. For the prediction of the Zika virus, a variety of classifiers are tried; eventually, the MLP classifier outperforms the others in terms of accuracy. The

use of encryption methods such as symmetric and asymmetric cryptography is investigated for the purpose of data security. For the purpose of encrypting the data, many methods are used. Finally, the suggested Hybrid Encryption technique is used for the purpose of safeguarding the data in the shortest amount of time possible, resulting in improved performance. The paper has the following structure: Section 2 consists of literature survey, Section 3 consists of methodology and outcome of proposed algorithm, and Section 5 consists of conclusions with future work.

2. Literature Survey

Data integrity is discussed in detail by Wang et al. (2015) [8], who propose that when a service provider provides many services to cloud users and those users share data in a group, the originality of the data should be maintained, and this is accomplished through public auditing, in which case the signature of the shared data blocks must match the signature of the service provider. Different blocks are signed by different users at multiple times when different users modify the same file. For security reasons, any user may have their access terminated before they have the opportunity to be resigned by an existing user. It is necessary to invoke the idea of public auditing in conjunction with an efficient user revocation procedure; proxy resigns are carried out by the cloud on behalf of the current user at the moment of revocation. Existing users will not be required to download and resign their licenses. Chen et al. (2012) [9] have claimed that cloud computing offers several benefits, such as the ability to host applications and data on the client's behalf and that clients and users are increasingly using hybrid or public clouds. The scale of their market prevents certain huge organizations and corporations from shifting data to the cloud for some mission-critical applications, and this is a problem. Users' recommendations and perspectives on security and privacy protection concerns are taken into consideration, and appropriate measures are taken.

The authors provide a comprehensive analysis of data security and protection in relation to privacy regulations across the whole life cycle of data that is discussed in more detail. Some of the existing solutions and research efforts connected to privacy and security concerns are presented, as well as some of the challenges that remain. Lopez-Barbosa et al. (2016) [10] made a proposal about the real-time utilization of Internet of Things devices, such as smart phones. Using sensor devices and cloud computing, Quwaider and Jararweh (2016) [11] have presented a method for increasing public health-related awareness in the community. It is necessary to utilize the map reduction idea in order to detect anomalies in the information supplied by the sensors in real time. Mamun and colleagues (2017) [12] detailed how a notion of delivering speech signals to a doctor using cloud technology may be implemented. The doctor then diagnoses the patient and keeps track of them with the use of mobile phones and the cloud, which is the suggested methodology's manner of operation. With the use of smart phone technology, Zhang et al. [2015] [13] suggested a method of monitoring and managing the epidemic. Based on their network contacts, the whole population is divided into several clusters, with outbreak methodologies being employed and deployed at the cluster level in this case. Sareen, Sood, and Gupta (2016) [14] made a proposal for intrusive technical enhancement in the Internet of Things, mobile computing, among other things. In parallel with enhancing the quality of service provided by technology, healthcare services are also being enhanced. Because of this virus, people have been affected in a variety of geographical locations. As a result, a wide range of neurological symptoms and infections were discovered and documented. As a result, care should be made to avoid contracting Zika since it is very contagious among pregnant women, babies, and adults alike. The existence of the Zika virus in India was established by Sumit Bhardwaj and colleagues (2017) [15]. On the basis of the samples, four of the Zika-infected patients were identified during the screening process. As a result, it may become a significant concern in the future. When compared to Chikungunya, Zika is expected to become a major concern in the near future.

The Zika epidemic, as represented by Petersen et al (2016) [16], as well as the infection among pregnant women and newborns, was addressed. Children with microcephaly are at a higher risk of developing neurological problems. A number of syndromes associated with neurological disorders were explored. This procedure has a high degree of accuracy, and the backpropagation approach was employed to anticipate the most accurate outcome possible. Kadri et al (2016) [17] have recommended that the Zika virus be designated a worldwide public health emergency. As a result, nations with a low risk of contracting the Zika virus were supplied with information pictorial expertise. It is necessary to take preventative measures.

It was suggested by Orellana et al. (2010) [18] that Google Docs have a new transparent user layer, which is implemented in Firefox, that encodes the record before collecting it in the Google server, making it impossible to access a data without having the correct password. The user is given the opportunity to choose the algorithms that will be used to encode the information. Once an algorithm has been chosen, the data is converted into cypher text and stored on Google's servers. The results demonstrate that blowfish performs much better when the key size is decreased, and the speed is increased. According to Singh et al. (2012) [19], elliptical curve cryptography is a fantastic approach of encryption technology. In wireless communication, the security layer must be implemented as robustly as possible, and for this reason, the framework is created using the ECC technique, which encrypts data in a powerful manner. It facilitates communication by using a multiagent system. The ECC has been used for the communication of wireless apps as well as certain web-based applications.

Singh et al. (2016) [20] have suggested a hybrid framework that incorporates both symmetric and asymmetric methods. When ECC and Blowfish are used together, the security level is significantly increased. The crypto and the CSP agent are both accessible for the purpose of distributing the key to the user. Even though CSP is not aware of it, the crypto agent is encrypting the data on its behalf.

Only the authorized individual will be able to decipher the information. User's private key is shared with CSP, and CSP's public key is shared with the user. As a result, not even the CSP was able to decode the data. CA is in charge of providing these services. Nathiya et al. (2019) [21] provided an explanation of the many network assaults that might occur when a packet is being sent. The intrusion detection approach is explained here and divided into four stages, with the goal of detecting attacks on cloud data storage as the data is being transferred. When an attack is introduced into the network, it is detected using false alarm methods, and the suggested algorithm HINDS is used to identify the attackers who have done so. When it comes to cloud storage security, Gampala et al. (2012) [22] introduced the ECC method as well as a digital signature mechanism for protecting information. With the use of the ECC technique, the security of the data is enhanced while the key size is reduced.

Jana et al. [23], in this case, used the hybrid approach, in which the downloading and uploading of data are completed at both the sender and recipient ends of the transmission. If any data is lost, it is impossible for both parties to decode it, which increases the overall security of the system. The multilayer algorithm is secure on both the user's and the server's end. Mohamed et al. (2015) [24] advocated that a framework be developed and validated for cloud environments in order to ensure that they are safe on both the client and server sides. When encoding or decoding data for connection setup, the Diffie Hellman cryptography is utilized in conjunction with ECC cryptography, and the integrity of the data is confirmed using the MD5 algorithm when updating the data. The suggested solution is a mix of ECC and SHA in order to provide a better outcome in data security. From the publications mentioned above, we can conclude that ECC is an improved asymmetric approach with a smaller key size. For the purpose of ensuring data integrity and authorization, a variety of encryption techniques and approaches were used. A brief overview of cloud computing is provided in the next chapter, followed by a detailed discussion of data security challenges and solutions for safeguarding data stored in the cloud.

3. Proposed Zika Virus Prediction Using MLP Classifier

Listed below is a description of each of the four components of the proposed system. The data gathering, fog layer, cloud layer, and, finally, the process are all in constant contact with the individuals engaged in the provision of healthcare. A framework has been designed for the identification of the Zika virus as well as the fortification of data in order to combat the virus. Cloud computing is used in this instance, because it is critical in the fact that it is capable of handling massive volumes of data from sensors and portable devices that have been mixed up. In order to connect end users to large-scale cloud services for storing and processing data, as well as for offering application services, it is essential to have a secure connection. Figure 3 depicts the whole architecture of the predicted model for predicting the Zika virus, which includes all of its components.

3.1. Input Data Generation. Synthetic data is information that has been manufactured artificially rather than via real-world data collection. In algorithmic testing, it is used to evaluate a dataset of operational data or a dataset from a production environment. Additionally, it may be used to the validation of mathematical codes and, to a greater extent, to the training of machine learning forms. It is used in the modelling of a scenario or the calculation of a theoretical value, among other things. It delivers an unexpected outcome and if the findings are found to be unsuitable, it gives the required cures or answers to the problem. The actual and confidential data for the basic test are replaced with synthetic data created by the test engine. It is occasionally necessary to produce synthetic data in order to safeguard the confidentiality of the concerned data. We are utilizing synthetic data to test all of the real-time events that occur. We were unable to get real-time data because we needed to protect the anonymity of the patients who had been afflicted with the Zika virus; therefore, we developed synthetic data. Our technique is an early prediction system, and we are able to forecast whether a patient is infected or not based just on the symptoms that they exhibit. Even with that, we were unable to get real-time information. It is similar to a real-time dataset in that we create the information of the patients, such as how many days he has been sick with fever and whether he is travelling to a high-risk location. The symptoms of the Zika virus are thoroughly examined in order to develop the suggested technique for conducting the tests. Because it is difficult to get patient information in India, we want to employ synthetic data, which will allow us to test all of the possible combinations based on our assumptions. In this case, the diagnosis of infection is made based on seven separate symptoms. The potential combinations of Zika virus symptoms samples are included in the following. Then, using synthetic data, the location of mosquito breeding sites and the location of mosquito dense sites are determined. As a result, the mapping is done randomly with respect to the area, symptoms, and the user. Therefore, it is simple to distinguish between infected and uninfected patients, as well as the preventative actions that should be implemented by government agencies and hospital personnel [25].

3.2. Input Data Tuning Layer. The suggested model includes a data component that comprises the specifics of the user's health data, environmental data, and location data, among other things. It is possible to get information on environmental conditions such as humidity, carbon dioxide level, and meteorological conditions by using environmental data. Because it is the primary cause of mosquito reproduction, it should be stressed repeatedly. Knowing well that our climatic conditions are ideal for mosquito reproduction, there is no need to watch everything every second. It is instead highlighted that the general climatic situation is favorable. The next step is to collect user health information. To do so, each user must register with the system using the mobile application that is available. Each user was assigned ID, which was generated. The indications of the Zika virus are obtained from users on a regular basis and reported to the

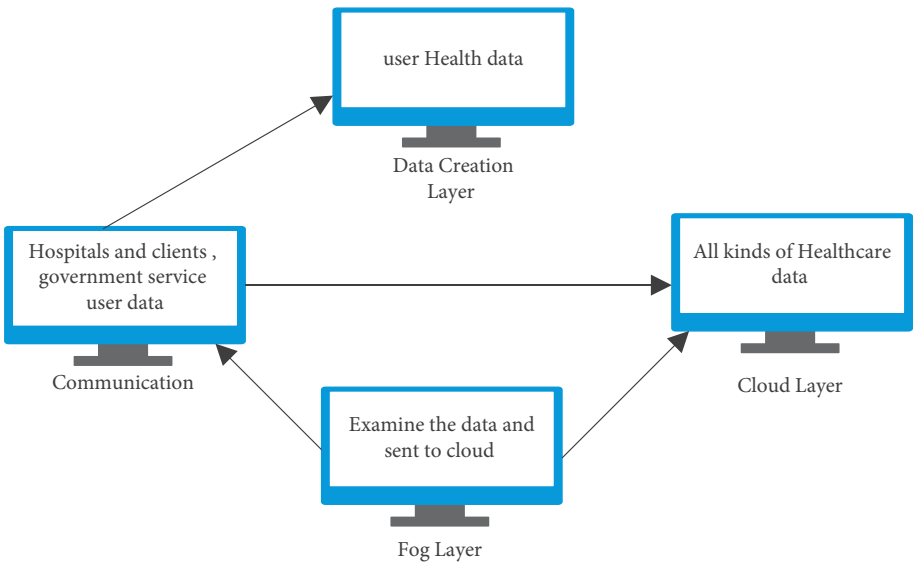


Figure 3: Proposed model of Secured Zika Virus Prediction.

authorities. The symptoms are responded in a yes or no fashion, according to the yes or no pattern. Not only are the symptoms recorded, but also is the user's health-related information. These kinds of information are gathered with the assistance of the sensor that is made available to the user [26]. The acquired data is protected using some kind of encryption technology, which is done in order of priority. User input is required for the symptoms of the Zika virus to be collected over a period of time, and the data are submitted as a yes or no pattern. The data collected by the environmental sensor includes information on mosquito breeding and population density. The sensor collects data in real time and uses it to pinpoint the location of the breeding grounds. In addition, carbon dioxide levels are regularly measured and studied in order to determine the climatic state of a certain site. Every piece of information pertaining to the environment is gathered in this section and saved in the fog computing servers. The data in the location part are connected to the data in the preceding section in that it displays the ideal site where there is a probability of mosquito density, and education opportunities are in height when compared to the climatic conditions. Table 1 represents the attributes of the input used in the proposed work.

Table 2 gives the prediction rate based on the symptoms, and Table 3 represents prediction rate based on the environmental criteria.

Table 1: Input attributes collected to predict disease.

S.NO	Input	Description
1	S.No	Serial number of input user
2	Reg ID	Registration id of the user
3	Gender	Gender of the user
4	Name	Name of the input user
5	Location	Primary address of the patient
6	Contact no.	Emergency contact number

Table 2: Prediction criteria based on symptoms.

S.NO	Input attributes	Inputs
1	High fever	N/y
2	Conjunctivitis	N/Y
3	More joint pain	N/Y
4	Allergic reaction	N/y
5	Inner muscular pain	N/y
6	Headache, vomiting	N/Y
7	Overall risk criteria	

Table 3: Prediction criteria based on environmental hazards.

Attributes	Narration
Dense areas mosquito available	Value obtained from GPS location
Input breeding area	Value obtained from dense breed area
Humidity	Stagnant temperature
Temperature	Normal
Carbon di oxide	Higher level humidity

3.3. Input Fog Computing Layer. Fog computing is a distributed computing environment that is used to handle large amounts of data in real time. It works as a platform between the cloud service provider and the user, allowing for large-scale data storage in the cloud to be accomplished. It is necessary to employ fog computing in order to reduce processing and performance time [27]. When it gathers all of the sensor data and stores it in a fog server, only the data that has been determined as necessary is evaluated and sent to the cloud for further processing. The processing speed and time are shortened as a result [28]. Because of the fog, the latency range, bandwidth, and everything else has risen. So, it serves as an independent server for data processing and archiving purposes. In the proposed work, fog is tasked with the responsibility of gathering all sensitive information from the user and determining if the symptoms match those of the

user. As a consequence, this sort of result is solely sent to the cloud. Fog is a first-level environment in which sensitive data acquired from the sensor must be kept in huge quantities due to the nature of the environment. As a result, there is a need to analyze the data and ensure that they correspond to the given one. This is followed by sending the data to the cloud, where the final data categorization and subsequent processing will take place.

3.4. Data Security. The acquired data are safeguarded via the use of a secret sharing method, in which the data are divided into little pieces and prioritizing is given for the various tiers. The level of protection provided for a piece of data is determined by its sensitivity. The protection of user personal data, which should be kept safe from the hands of unauthorized individuals, is given the highest priority. In the second level, there is information about the environment, and the information is saved on multiple servers. The third item is the least important since it should include information on symptoms as well as a warning to the individual to take the required precautions to prevent contracting the Zika virus. The hospitals in the government-run healthcare system provide the essential guidance to the public. Figure 4 represents the overall proposed system.

3.5. Classification Using Multilayer Perceptron with Probability Optimization. The probabilistic model-based classifier is based on the mean and variance measures of the produced classes, which are a total of 124 classes in this case (62 classes of the NN and 62 classes of DT). The mean and variance measurements are used to calculate the exterior probability value of the approved picture, which is then expressed as a percentage. In order to avoid overfitting, we estimate the exterior probability of the 73 classes that are important to both the NN and the DT classifiers independently. The resulting probability of the NN classifier is multiplied by the resultant probability of the DT classifier to generate the new probability value for the classifier. The most exact recognition of the character picture is obtained in line with the greatest value of the new posterior probability distribution. The procedures involved in the modelling of the probabilistic classifier for the recognition of character are explored in further detail in the following sections.

In the probability calculation using (1),

$$(G \times 2) = \left[D_0 \;\; E_f \right]; \quad 1 \le o \le 62. \tag{1}$$

In this case, o is the output class label, and the class label of the NN classifier is indicated as Do, whereas the class label of the DT classifier is written as E. It is possible to define the mean and variance metrics associated with the NN and DT classifiers in the form of equations (2) and (3):

$$\text{Mean value } = o \times 2 \left[N(D_o) \;\; N(E_o) \right], \tag{2}$$

$$\text{Variance value } = o \times 2 \left[W(D_o) \;\; W(E_o) \right]. \tag{3}$$

The mean value is denoted by the letter N0, and the variance value is denoted by the letter W 0 in this equation. The 124 mean values are represented as o2, where 62 mean values correspond to the mean value of the NN classifier and the remaining 62 mean values correspond to the mean value of the DT classifier. Furthermore, the variance values of the NN and DT classifiers are computed and compared.

NN classifier and DT classifier posterior probability formulas are shown in the following table: NN classifier posterior probability formula.

According to the following equation, the posterior probability formula for the NN classifier may be found:

$$\mathbf{Q}\left(\frac{\mathbf{J}}{\mathbf{D}_b}\right) = \left[\frac{f\left[-\left((\mathbf{D}_\mathrm{b} - \mathbf{N}(\mathbf{D}_\mathrm{B}))^2 / (2 + \mathbf{W}(\mathbf{D}_\mathrm{b}))^2 \right)\right]}{\sqrt{2 * \pi * W(\mathbf{D}_\mathrm{b})}} \right]. \tag{4}$$

NN class label mean value is represented as M (C s), and NN class label variance measure is represented as W (D s) in this example.

According to the following equation, the posterior probability formula for the DT classifier may be found:

$$\mathbf{Q}\left(\frac{\mathbf{J}}{\mathbf{E}_b}\right) = \left[\frac{\left[-\left((\mathbf{E}_\mathrm{f} - \mathbf{N}(\mathbf{E}_\mathrm{f}))^2 / (2 * \mathrm{W}(\mathbf{E}_\mathrm{f}))^2 \right)\right]}{\sqrt{2 * \pi * \mathrm{W}(\mathbf{E}_\mathrm{g})}} \right]. \tag{5}$$

The mean value of the DT class label is denoted by the letter N (E o), while the variance measure of the DT class label is denoted by the letter W (E o).

NN classifiers have posterior probabilities that are higher than the posterior probability of the DT classifiers, which is represented by the formula for maximum posterior probability (max posterior probability) (5). The probabilistic model, as shown in equation (6), is used to identify the input character picture with the highest likelihood of being recognized:

$$D = \max_{b=1}^{62} |Q(I \;\; D)|. \tag{6}$$

On the basis of the greatest measure of posterior probability, the identification of a character is conducted. In line with the acceptance criteria, the optical character picture with the highest likelihood of being identified properly is selected from the class.

Figure 5 represents the Multilayer Perceptron Neural Network. Each of the NN and DT classifiers has its own mean value and variance value, which are computed independently in the algorithm. The posterior probability value is then computed using the method given above, which takes into account the mean and variance values that have been acquired. The posterior probability values of the NN and DT classifiers are merged to generate a single probability value, which is then blended with the other posterior probability values. Final recognition is achieved by using the greatest probability measure to determine whether or not the input character is recognized.

For identifying the instances, the multilayer perceptron classifier employs the backpropagation approach, which is described in more detail below. The network was built via the MLP algorithm, which was then analyzed and tweaked throughout the training phase. Except for the numeric classes, the network is composed entirely of sigmoid nodes

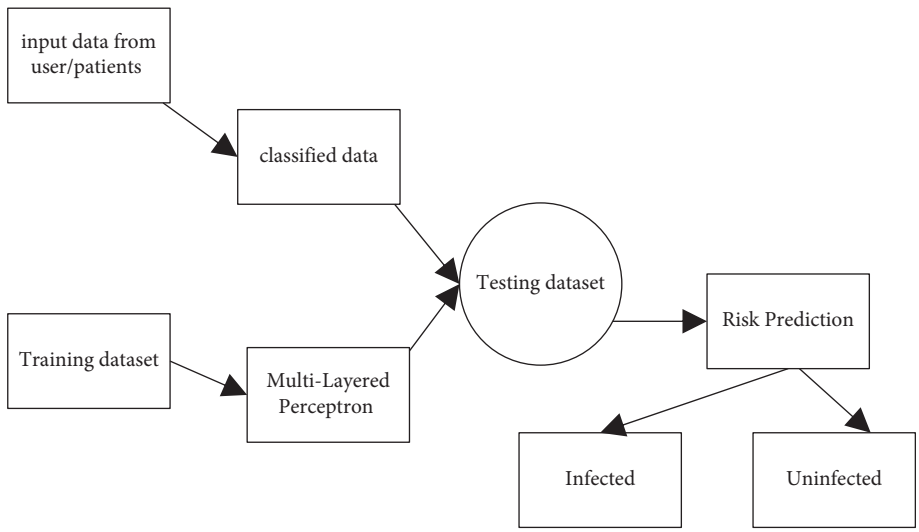


Figure 4: Overall flow of Proposed work using MLP Classifier.

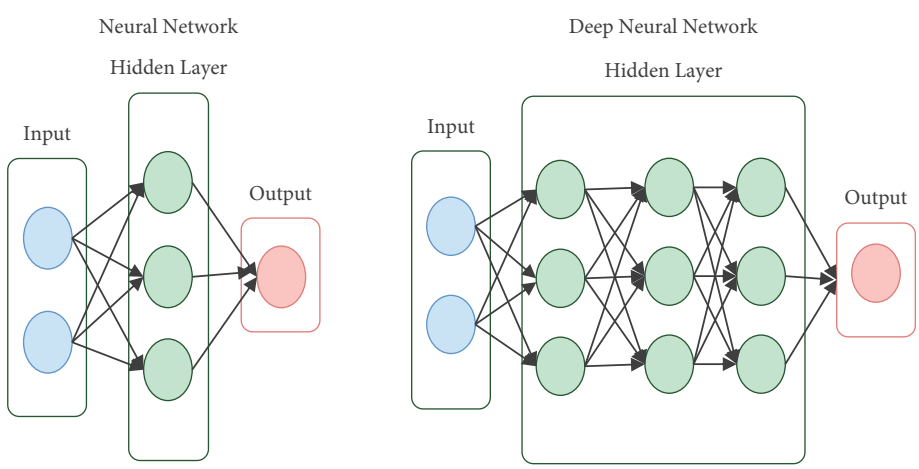


Figure 5: Multilayer perceptron neural network. (a) Neural Network. (b) Deep neural network.

for the threshold value. The backpropagation technique is required for the elements in order to get a complicated output result. It operates on the inputs in the network using the feedforward mechanism. When performing the iterative method, a set of weights is used to forecast the class label for each iteration. The feedforward algorithm contains a single input layer, one or more hidden layers, and eventually a single output layer, as shown in the diagram. During the classification process, if a mistake happens, the back-propagation approach is used to enhance the accuracy of classification while simultaneously decreasing the amount of input values and the time required for training.

$$\begin{aligned} \text{layer}_l &= \sum y_{kl} z_l + \theta_2, \\ P_k &= y_l - \frac{1}{1 + e^{-Ofl_k}}. \end{aligned} \tag{7}$$

It is computed by feeding the production spinal into the hidden film, which then processes the contribution over using the feedforwarding method, which results in a sigmoid function.

3.6. Weight and Objective Function of Classifier. The output node k has an activation value of and an anticipated target value of tk for node k, and the change among the predictable and authentic target values is represented as

$$B_4 - B2 - 0, \tag{8}$$

and node k is defined as

$$\begin{aligned} P_2 &= \hat{P}_2 Q_2 (M - P_2), \\ \Delta w_{jk} &= 4, \delta_L x_6, \\ w_{jL} &= w_{ja} + \Delta w_{jw}. \end{aligned} \tag{9}$$

The network is recognized based on the pace at which it is learning. If it is set too low, network learning will be very sluggish, and if it is set too high, the network will oscillate between minimum and maximum values. Altering the knowledge rate from a big to a minor value during the backpropagation technique has a number of benefits. Assume that a network begins with weights that are far from the set of optimal weights and that it receives rapid training

initially. When the learning rate falls throughout the course of learning, it is claimed that the process has reached a minimal optimum point. Because overshooting is less likely when the learning processes slow down. Figure 6 shows the hybrid structure of AES with DES.

The suggested hybrid technique includes two levels of encryption, or two tiers of security, and it is a framework that supports both symmetric encryption and asymmetric encryption functions at the same time. The two algorithms ECC and AES are a mix of hybrid methods for safeguarding data that are stored in the Common Security Policy (CSP). It is necessary to employ the AES symmetric method at the first level of encryption since the key is known to the user. The output of the first level of encryption is then encrypted a second time using the ECC algorithm, which is an asymmetric technique of encrypting data. In this case, the AES key is shared with both the user and the CSP in the first level, whereas in END USER Cloud Database Encryption using AES Level1 Encryption/Decryption using ECC Level 2 Data Collection the second level, the public key is used by the CSP and the private key is generated and shared with the user alone in the third level of the encryption scheme. The research of hybrid encryption method with two levels of encryption technology that went awry is shown in the illustration below. Each user will be provided with a set of two keys.

$$\begin{aligned} &\text{Key } z = \text{AESj}, \\ &\text{Triple DES } Q_J. \end{aligned} \tag{10}$$

In this case, AESi represents the ith user of the symmetric key, and it is only known to the user. ECCpri displays the ith user's asymmetric private key, which is solely used by that particular user and no one else. The asymmetric public key user I is represented by the ECCpui, and it is known to the CSP. When a user saves data in the cloud, the cloud service provider (CSP) provides a set of keys that are used to encrypt the data. First, the data is encrypted using a key that is only known to the user and no one else. Before it is stored in the cloud storage, the cypher text generated as a consequence of the calculation is encrypted once again. In the next step, the completed encrypted text is saved in cloud storage [20–22]. Even the CSP is unaware of the key required to decode the data. In this case, the user will only be responsible for the first level of decryption. The key for second-level decryption is known to the user in this case. This process of delivering service to both the user and the CSP is carried out with the assistance of a third-party agent known as the Crypto Provider.

Provider of cryptographic services: client-side encryption and decryption are handled by the client-side cryptographic processor (CP) [29]. It is ready after the set of users' keys has been collected. Whenever a CSP registers a new user, the CP makes the ECCpui available to the CSP. If a user wishes to keep data in cloud storage, he or she must first encrypt the data using the AESi algorithm and then using the ECCpui algorithm.

Pseudocode of Proposed work.

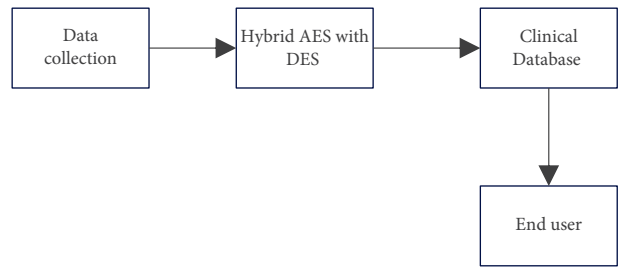


Figure 6: Hybrid AES with triple DES.

Preparation of the references no and S (i);

If (Refno = n), then the record already exists; else, the record does not exist.

Alternately, update the user's information.

Create a new entry in the database and save it there;

If refno (updated data = old data) is true, then the values should be stored in the existing data.

Else

Keep track of the categorization value in the database;

Change the patient's classification.

In cases where S i corresponds to a i, where $i \geq n$ (number of patients), class C is equal to IN; otherwise, class C is equal to UIN.

Inform the patient's infection classification;

An alert has been sent to the doctor, the patient, and the hospitals.

If the condition is met, the if statement is terminated.

4. Experimental Results

In order to make predictions about the Zika virus, the Matlab version is employed. It is a data mining software package that comprises a variety of machine learning algorithms for processing large amounts of data. It analyzes data in the ARFF file format by default, with CSV as a supported file format for Weka as a second option [30]. The explorer option is utilized for both the training set data and the testing dataset. The synthetic data was constructed using all feasible combinations of our assumptions and assumptions from other sources. Because it is regarded to be an uncommon illness in India, it is difficult to gather information about it.

Figure 7 represents the encrypted data. The MATLAB tool makes use of the dataset that was supplied in the previous section. It is necessary to utilize all of the possible combinations of the symptoms. Then, there are approximately 500 instances and 15 characteristics added to the mix. The individual is classed as infected or uninfected based on the information they provide. In this case, the MLP classifier is utilized to categorize the instances into groups. The MLP classifier has a 97 percent accuracy rate, which is excellent.

The occurrences are classed as infected or uninfected based on the true positive and false positive rates for each case. The dataset under this study is publicly available. The

Enter plain text: Eid al-Fitr 2020

Enter key: Sunday, May 24

CIPHER TEXT

['45', '61', '69', '32']

['69', '6c', '74', '30']

['64', '2d', '72', '32']

['20', '46', '20', '30']

CIPHER KEY

['53', '61', '4d', '32']

['75', '79', '61', '34']

['6e', '2c', '79', '0']

['64', '20', '20', '0']

ENCRYPTED TEXT

['0×c3', '0×23', '0×dc', '0×82']

['0×c', '0×62', '0×bc', '0×e6']

['0×9d', '0×fe', '0×dd', '0×b0']

['0×8', '0×ed', '0×53', '0×2d']

ENCRYPTED TEXT

Ã#bþíÜ¼ÝS,æ°-

Figure 7: Encrypted data.

algorithm for determining the sensitivity and specificity of the cases is used to determine these characteristics.

$$\text{Sensitivity} = \frac{\text{True Positive}}{(\text{True Positive} + \text{False Negative})},$$
$$\text{Accuracy} = \frac{(\text{True Negative} + \text{True Positive})}{(\text{Tue Negative} + \text{True Positive} + \text{False Negative} + \text{False Positive})}. \tag{11}$$

Figure 8 represents the decrypted data. The threshold value for each class probability ranges from one class to the next. As a result, a classifier that yields an MLP threshold is described. It is indicated on the X-axis that the example dimensions are supplied and on the Y-axis that the true positive rate is presented.

Figure 9 represents the classification accuracy. Let us say that the sensitivity, specificity, and accuracy of the aforementioned equations are determined. It is determined whether or not the diagnostic test is accurate based on these data. Following that, the specificity of the test indicates the usual diagnostic situation, which is a negative result. Table 4 provides an overview of the comparative examination of several classification algorithms.

While accuracy is defined as the ability to accurately detect the genuine result of the whole population, it is concerned with the real severity of a test diagnostic condition. A tall compassion test is designed to capture all of the potential positive situations that might occur during a test. As a result, sensitivity is utilized in the screening of diseases. When compared to other mosquito-borne diseases, the symptoms are mild to moderate. There is an incubation period of around 5 days for the virus [32]. If the symptoms linger for more than 7days, the individual should consult for an early prediction approach that employs our suggested method. If they are confirmed to be contaminated, they must have an RNA test in which their RNA is thoroughly examined and analyzed.

DECRYPTED TEXT

['0×45', '0×61', '0×69', '0×32']

['0×69', '0×6c', '0×74', '0×30']

['0×64', '0×2d', '0×72', '0×32']

['0×20', '0×46', '0×20', '0×30']

DECRYPTED TEXT (Plain)

Eid al-Fitr 2020

KEY (Plain)

Sunday, May 24

Figure 8: Decrypted data.

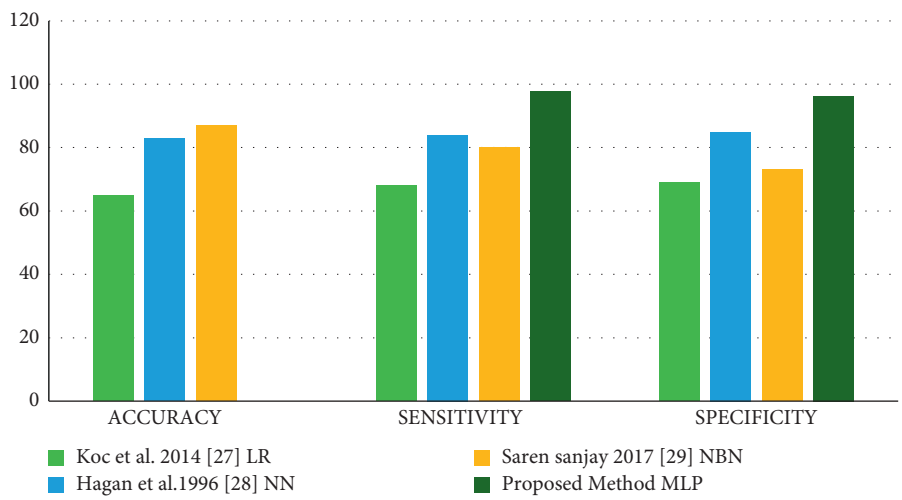


Figure 9: Classification accuracy.

Table 4: Input classification table.

Author	Methods	Accuracy	Sensitivity	Specificity
Koc et al. 2014 [27]	LR	65	68	69
Hagan et al.1996 [28]	NN	83	84	85
Saren Sanjay 2017 [31]	NBN	87	80	73
Proposed method	MLP	97.5	97.63	96.28

5. Conclusion

To determine whether or not the user is infected, the suggested system is utilized to gather data from the user and, depending on the symptoms, diagnose the user using the MLPNN algorithm for improved accuracy. The common characteristics of environmental factors are ready to be used to create a risky environment that is susceptible to infection. Once infections in patients have been found, the information pertaining to those diseases must be safeguarded in cloud storage. In order to safeguard the data stored in cloud storage, a double layer encryption approach using a hybrid crypto algorithm is used. The data is encrypted using the ECC and AES encryption methods, and even the third-party supplier has little knowledge of the contents of the encrypted data. Our suggested approach makes use of categorization to provide better results with an accuracy of 98 percent, and it assists the government's primary healthcare department in controlling the number of mosquitoes that are reproducing extremely successfully in the area. We are able to provide a better solution for the Zika virus infection when the healthcare industry and the government work together to implement our technology. The increased accuracy achieved in this research will be adopted, which means that it will assist physicians in the accurate prediction of Zika virus and the reduction of microcephaly illness in newborns and fetuses, among other things. Even premature delivery was averted to some extent. Patients in India must be monitored,

and if any of the symptoms listed above are observed in any of them, the prediction system will take care of the prediction, as well as data protection, which is extremely secure, and the healthcare sectors do not essential to be concerned about the information being kept in fog storage because of the HEA used in the proposed model, which stands for Health Equity Act. Using the prediction system for anything beneficial to human civilization is the long-term goal of the project. The RNA test is the second step in the Zika virus prediction process. It is intended to concentrate on prevention and preventing the spread of the Zika virus. When it comes to cloud computing, new technologies are being developed on a daily basis, and data breaches are also occurring, so we must be prepared to deal with both the ups and downs. Because of this, research should be conducted for the benefit of society as a whole.

Data Availability

The data that support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors of this manuscript declare that they do not have any conflicts of interest.

References

[1] S. M. Kamal and K. El Sayed Khalifa, "Immune modulation by helminthic infections: worms and viral infections," *Parasite Immunology*, vol. 28, no. 10, pp. 483–496, 2006.

[2] K. Kumar, A. K. Sharma, M. Sarkar, A. Chauhan, and R. Sharma, "Surveillance of *Aedes aegypti* (L.) mosquitoes in Mumbai international seaport (India) to monitor potential global health risks," *Journal of Insects*, vol. 2014, Article ID 951015, 5 pages, 2014.

[3] W. Diffie and M. Hellman, "New directions in cryptography," *IEEE Transactions on Information Theory*, vol. 22, no. 6, pp. 644–654, 1976.

[4] C. Alcaraz and S. Zeadally, "Critical control system protection in the 21st century," *Computer*, vol. 46, no. 10, pp. 74–83, 2013.

[5] M. El Ghazouani, M. A. El Kiram, and L. Er-Rajy, "Blockchain & multi-agent system: a new promising approach for cloud data integrity auditing with deduplication," *International Journal of Communication Networks and Information Security*, vol. 11, no. 1, pp. 175–184, 2019.

[6] E. Garbarino and M. Strahilevitz, "Gender differences in the perceived risk of buying online and the effects of receiving a site recommendation," *Journal of Business Research*, vol. 57, no. 7, pp. 768–775, 2004.

[7] J. Ryoo, S. Rizvi, W. Aiken, and J. Kissell, "Cloud security auditing: challenges and emerging approaches," *IEEE Security & Privacy*, vol. 12, no. 6, pp. 68–74, 2013.

[8] Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li, "Enabling public auditability and data dynamics for storage security in cloud computing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 22, no. 5, pp. 847–859, 2010.

[9] S. Luo, Z. Lin, X. Chen, Z. Yang, and J. Chen, "Virtualization security for cloud computing service," in *Proceedings of the 2011 International Conference on Cloud and Service Computing*, pp. 174–179, IEEE, Hong Kong, China, 2011 December.

[10] N. Lopez-Barbosa, J. D. Gamarra, and J. F. Osma, "The future point-of-care detection of disease and its data capture and handling," *Analytical and Bioanalytical Chemistry*, vol. 408, no. 11, pp. 2827–2837, 2016.

[11] M. Quwaider and Y. Jararweh, "Cloudlet-based efficient data collection in wireless body area networks," *Simulation Modelling Practice and Theory*, vol. 50, pp. 57–71, 2015.

[12] S. Sondhi, S. Saad, K. Shi, M. Mamun, and I. Traore, "Chaos engineering for understanding consensus algorithms performance in permissioned blockchains," 2021, https://arxiv.org/abs/2108.08441.

[13] N. H. Yu, Z. Hao, J. J. Xu, W. M. Zhang, and C. Zhang, "Review of cloud computing security," *Acta Electonica Sinica*, vol. 41, no. 2, p. 371, 2013.

[14] S. Sareen, S. K. Gupta, and S. K. Sood, "An intelligent and secure system for predicting and preventing Zika virus outbreak using Fog computing," *Enterprise Information Systems*, vol. 11, no. 9, pp. 1436–1456, 2017.

[15] P. Gupta, S. Bharadwaj, and V. K. Sharma, "A survey to bridging the gap between energy and security in IoT and home," in *Proceedings of the 2019 Fifth International Conference on Image Information Processing (ICIIP)*, pp. 379–384, IEEE, Shimla, India, 2019 November.

[16] W. A. Petersen and S. A. Rutledge, "On the relationship between cloud-to-ground lightning and convective rainfall," *Journal of Geophysical Research: Atmospheres*, vol. 103, no. D12, pp. 14025–14040, 1998.

[17] L. Benarous and B. Kadri, "Obfuscation-based location privacy-preserving scheme in cloud-enabled internet of vehicles," *Peer-to-Peer Networking and Applications*, pp. 1–12, 2021.

[18] D. A. Rodrıguez-Silva, F. J. González-Castano, L. Adkinson-Orellana, A. Fernández-Cordeiro, J. R. Troncoso-Pastoriza, and D. González-Martınez, "Encrypted Domain Processing for Cloud Privacy," in *Proceedings of the 1st International Conference on Cloud Computing and Services Science*, Noordwijkerhout, Netherlands, May 2011.

[19] S. Singh, Y.-S. Jeong, and J. H. Park, "A survey on cloud computing security: Issues, threats, and solutions," *Journal of Network and Computer Applications*, vol. 75, pp. 200–222, 2016.

[20] A. Singh, D. Juneja, and M. Malhotra, "Autonomous agent based load balancing algorithm in cloud computing," *Procedia Computer Science*, vol. 45, pp. 832–841, 2015.

[21] T. Nathiya and G. Suseendran, "An effective hybrid intrusion detection system for use in security monitoring in the virtual network layer of cloud computing technology," in *Data Management, Analytics and Innovation*, pp. 483–497, Springer, Singapore, 2019.

[22] V. Gampala, S. Inuganti, and S. Muppidi, "Data security in cloud computing with elliptic curve cryptography," *International Journal of Soft Computing and Engineering*, vol. 2, no. 3, pp. 138–141, 2012.

[23] B. Jana, M. Chakraborty, and T. Mandal, "A task scheduling technique based on particle swarm optimization algorithm in cloud environment," in *Soft Computing: Theories and Applications*, pp. 525–536, Springer, Singapore, 2019.

[24] E. M. Mohamed, H. S. Abdelkader, and S. El-Etriby, "Enhanced data security model for cloud computing," in *Proceedings of the 2012 8th International Conference on Informatics and Systems (INFOS)*, IEEE, Giza, Egypt, 2012 May.

[25] X. Li, Y. Zheng, M. D. Alshehri et al., "Cognitive AmBC-noma IoV-MTS networks with IQI: reliability and security analysis," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12, 2021.

[26] L. M. Anderson, S. C. Scrimshaw, M. T. Fullilove, J. E. Fielding, J. Normand, and Task Force on Community Preventive Services, "Culturally competent healthcare systems," *American Journal of Preventive Medicine*, vol. 24, no. 3, pp. 68–79, 2003.

[27] D. Samanta, M. P. Karthikeyan, D. Karuppiah et al., "Optimized tree strategy with principal component analysis using feature selection-based classification for newborn infant's Jaundice symptoms," *Journal of Healthcare Engineering*, vol. 2021, Article ID 9806011, 9 pages, 2021.

[28] A. S. Rajawat, P. Bedi, S. B. Goyal et al., "Fog big data analysis for IoT sensor application using fusion deep learning," *Mathematical Problems in Engineering*, Article ID 6876688, 16 pages, 2021.

[29] B. L. Dey, A. Pandit, M. Saren, S. Bhowmick, and H. Woodruffe-Burton, "Co-creation of value at the bottom of the pyramid: analysing Bangladeshi farmers' use of mobile telephony," *Journal of Retailing and Consumer Services*, vol. 29, pp. 40–48, 2016.

[30] X. Li∗, M. Zhao, M. Zeng et al., "Hardware iab systems: reliability and security," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2723–2736, 2021.

[31] M. L. Koç and D. I. Koç, "A cloud theory based reliability analysis method and its application to reliability problems of breakwaters," *Ocean Engineering*, vol. 209, Article ID 107534, 2020.

[32] M. Hagan, F. Siddiqui, S. Sezer, B. Kang, and K. McLaughlin, "Enforcing policy-based security models for embedded SoCs within the internet of things," in *Proceedings of the 2018 IEEE Conference on Dependable and Secure Computing (DSC)*, pp. 1–8, IEEE, Kaohsiung, Taiwan, 2018 December.

Hindawi
Journal of Healthcare Engineering
Volume 2022, Article ID 4182191, 11 pages
https://doi.org/10.1155/2022/4182191

Research Article

CXR-RefineDet: Single-Shot Refinement Neural Network for Chest X-Ray Radiograph Based on Multiple Lesions Detection

Cong Lin,[1] **Yongbin Zheng**,[1] **Xiuchun Xiao**,[1] **and Jialun Lin**[2]

[1]*College of Electronics and Information Engineering, Guangdong Ocean University, Zhanjiang 524025, China*
[2]*College of Biomedical Information and Engineering, Hainan Medical University, Haikou 571199, China*

Correspondence should be addressed to Xiuchun Xiao; xiaoxc@gdou.edu.cn

Received 24 October 2021; Revised 3 December 2021; Accepted 6 December 2021; Published 7 January 2022

Academic Editor: Han Wang

The workload of radiologists has dramatically increased in the context of the COVID-19 pandemic, causing misdiagnosis and missed diagnosis of diseases. The use of artificial intelligence technology can assist doctors in locating and identifying lesions in medical images. In order to improve the accuracy of disease diagnosis in medical imaging, we propose a lung disease detection neural network that is superior to the current mainstream object detection model in this paper. By combining the advantages of RepVGG block and Resblock in information fusion and information extraction, we design a backbone RRNet with few parameters and strong feature extraction capabilities. After that, we propose a structure called Information Reuse, which can solve the problem of low utilization of the original network output features by connecting the normalized features back to the network. Combining the network of RRNet and the improved RefineDet, we propose the overall network which was called CXR-RefineDet. Through a large number of experiments on the largest public lung chest radiograph detection dataset VinDr-CXR, it is found that the detection accuracy and inference speed of CXR-RefineDet have reached 0.1686 mAP and 6.8 fps, respectively, which is better than the two-stage object detection algorithm using a strong backbone like ResNet-50 and ResNet-101. In addition, the fast reasoning speed of CXR-RefineDet also provides the possibility for the actual implementation of the computer-aided diagnosis system.

1. Introduction

Chest X-ray (CXR) is an effective and widely used imaging technique in the diagnosis and screening of lung-related diseases. The imaging principle and structure of chest radiographs are complex, which requires professional radiologists to spend a lot of time to observe carefully. Medical research [1, 2] shows that postprocessing of medical images by using a computer-aided diagnosis (CAD) system can effectively reduce the initial screening of chest radiographs and improve the accuracy of lesion screening. Benefiting from the rapid development of the field of artificial intelligence [3], many researchers have proposed lots of automatic diagnosis methods by combining deep learning technology with imaging examination technology to reduce the workload of radiologists and the possibility of misdiagnosis [4, 5]. The use of deep learning technology to assist doctors in diagnosing diseases has become a new trend.

The application of deep learning in the field of medical imaging is mainly in the two major tasks of segmentation and detection. Because of the reason that the segmentation task can provide doctors with precise lesion area positions and is more suitable for actual needs, it has become a research hotspot in the field of medical imaging [6, 7]. Arnaud proposed a new computer-aided detection lung nodule system using multiview convolutional networks (ConvNets) to reduce the false positives of the CAD system [8]. Olaf proposed a segmentation network U-Net [9] that efficiently utilizes medical image annotation, which vigorously promotes the development of medical image segmentation. Rocha [10] proposed a lung nodule segmentation method based on U-Net and SegU-Net to solve the segmentation problem of lung nodules in computed tomography. The segmentation results of lung nodules by this method can help doctors further analyze the lesions feature. However, the segmentation model may be inaccurate for the

segmentation of the edge part of the lesion and the small lesion area in practical applications since the spatial dimension of a chest X-ray is usually 2000×3000 pixels and the local lesion area is relatively small, which makes the detection more difficult and requires the doctor to spend more time to make further judgments.

Different from image segmentation, object detection provides a candidate area of the lesion, which will help the doctor to quickly locate the lesion area instead of focusing on the pixel-level segmentation area [11]. In terms of lung disease screening, studies [12–15] have shown that detection models designed through deep learning can provide doctors with areas where lung disease may occur, which can greatly improve the efficiency of disease screening by radiologists. In order to detect the location, type, and attributes of lung lesions more accurately, Yan designed a deep learning module that extracts relevant semantic tags from radiology reports related to lesion images. Using image and text to mine tags, a lesion annotation network (LesaNet) based on multilabel convolutional neural network is proposed to learn all the tags in a comprehensive way [16]. Liang proposed a method to filter out target images with lung nodules from the patient's whole lung CT images by training a classification network and then use Faster RCNN to detect the location range of the suspected lung nodules in the CT image to increase the reliability of the detection and reduce the false positives of detectors [17]. Xiao [18] proposed a fully automatic lung nodule detection algorithm using a cascade strategy based on FPN [19]. By designing multiple stages of detection networks and integrating them into a heterogeneous classification network, the nodules are gradually separated from the lung background. Although the lung lesion detection methods based on the two-stage object detection model have higher detection accuracy, the two-stage object detection model is larger and occupies high computing resources, and it is difficult to deploy to the detection system of the hospital. Therefore, most of the current lesion detection models of medical images tend to use a combination of one-stage object detection algorithms and large backbone, such as using RetinaNet [20] as the base model and matching large backbone such as ResNet-101 [21] and SE-ResNet-101 [22] to detect lung lesions. Although this can increase the detection accuracy while speeding up the detection speed, these base models still require high computing resources and cannot be well applied in medical auxiliary detection systems.

From the above analysis, it can be seen that the high requirements of computing resources, slow reasoning speed, and low accuracy of detecting lesions hinder the application of artificial intelligence technology in the field of medical images. In order to solve the above problems and promote the implementation of computer-aided diagnosis technology in chest radiography, a chest radiograph lesion detection algorithm with a small model, high accuracy, and fast detection speed based on RefineDet network architecture is proposed in this paper. In the first step, we designed a backbone RRNet with a small amount of parameters and strong feature extraction capabilities by combining the advantages of RepVGG block and Resblock in information fusion and information extraction, which can improve the feature extraction capabilities of the network while reducing the amount of model calculations. Then, we proposed a structure called Information Reuse through connecting the normalized features of RefineDet back to the network again, which can effectively solve the problem of low utilization of the original network output features and achieve the purpose of improving detection accuracy. Combining the network of RRNet and the improved RefineDet, we propose the overall network named CXR-RefineDet. A large number of experiments have been done to verify the performance of CXR-RefineDet on VinDr-CXR [23], and the experimental results show that the RRNet backbone and Information Reuse structure we designed have brought about 0.99% and 0.72% improvement in detection performance, respectively. In addition, we also compare with the current mainstream object detection network on the three performance indicators of mAP, inference speed, and parameter amount. As shown in Figure 1, the comparative experiment results show that the detection accuracy and speed of the CXR-RefineDet network greatly exceed the existing mainstream object detectors under the condition of moderate parameter amount, which can effectively help doctors quickly and accurately screen the location of lesions in the image.

The main contributions of this work are as follows: (1) By combining the advantages of RepVGG block and Resblock in information extraction and information fusion, we designed a backbone RRNet with few parameters and strong feature extraction capabilities. (2) We propose the Information Reuse structure, which solves the problem of low utilization of the original network output features by linking the normalized features back to the network. (3) The proposed object detection model CXR-RefineDet has a good performance between accuracy and speed. It achieves 0.1686 mAP and 6.8 fps on the VinDr-CXR dataset, which is significantly better than mainstream object detection models.

2. Materials and Methods

2.1. Related Works. The one-stage object detection network RefineDet [24] proposed by Zhang adds Anchor Refinement Module (ARM) and Object Detection Module (ODM) to the network to perform preliminary filtering and further filtering of anchor frames, respectively. At the same time, the network also uses the Transfer Connection Block (TCB) module to fuse the features between ARM and ODM, so that the one-stage object detection network has the accuracy of the two-stage object detection network while maintaining a faster detection speed, which could make it possible for the model to be implemented. It should be pointed out that the backbone of RefineDet has two types, VGG-16 [25] and ResNet-101, but ResNet-101 has a large amount of parameters and requires high computing resources. Compared with ResNet-101 which stacks more residual blocks, VGGnet has the advantages of fewer network parameters and faster running speed. Considering that the efficiency of the medical auxiliary detection system is more important in practical applications, choosing VGG-16 as the backbone of

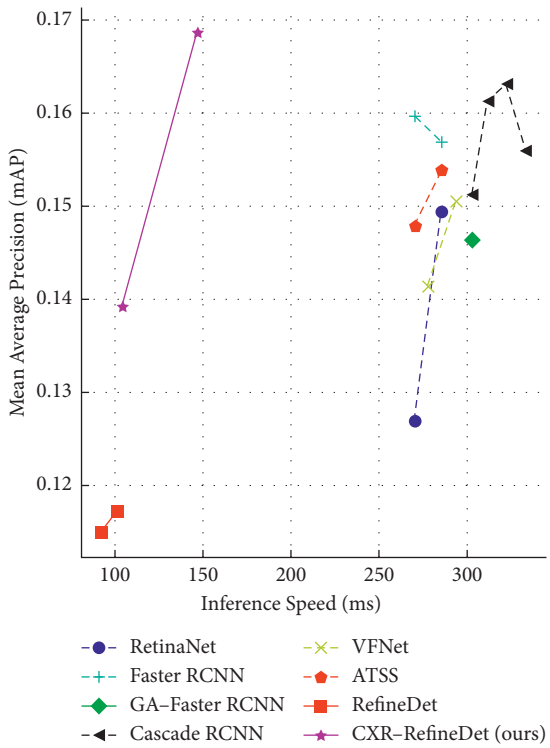


Figure 1: Comparison of the proposed CXR-RefineDet and other mainstream object detectors. CXR-RefineDet at 512 resolution reaches 0.1686 mAP, surpassing all mainstream detectors, and its single image inference speed is nearly 2 times that of other detectors.

RefineDet for medical image detection will be more practical. The contribution of VGG-16 proposed by Karen is a thorough evaluation of networks of increasing depth using an architecture with very small convolution filters, which shows that a significant improvement on the prior-art configurations can be achieved by pushing the depth to 16–19 weight layers [25]. Recently, many researchers have proposed excellent variant networks based on VGG. Huang [26] connected the low-level semantic information of the network with many high-level semantic information to build a complex network topology Densenet. Inspired by the ResNet and Inception structures, [27] designed the ResNext network structure by adding residual connections to the Inception structure. Zhang constructed the ResNeSt [28] network by introducing the Split Attention module and SKNet-block [29] on the basis of ResNeXt. Although new network structures are emerging one after another and the accuracy of the network has also been greatly improved, the amount of network parameters and requirements for computing resources have also become higher and higher. In addition, various new network module functions are complicated to implement, which further aggravates the difficulty of model deployment.

2.2. Network Architecture. In the previous analysis, it was mentioned that ResNet-101 needs to consume more computing resources, and the residual structure in ResNet requires that the feature dimensions before and after the residual must match, which limits the flexibility of the network, so we chose VGG-16 as the backbone of RefineDet. At the same time, we noticed that the RepVGG [30] network proposed by Ding is only composed of 3×3 convolution, BN layer, and ReLU modules, which is very beneficial to the acceleration of the neural network of mobile devices. In addition, since Resblock can fuse the feature information between multiple convolutional layers through jump connections between layers, and RepVGG block can improve the feature extraction ability of single-layer convolution after paralleling multiple convolution modules in a single convolution layer, we believe that combining the advantages of RepVGG block and Resblock in single-layer convolution and multilayer convolution can greatly improve the detection ability of the backbone. Considering that the network shallow convolution is responsible for extracting low-level semantic features, and the information richness of this part of the low-level semantic features directly determines the effectiveness of the high-level semantic features of the subsequent convolutional layer, we set the first three layers of the new backbone named RRnet as RepVGG block and use the feature of multiple modules in parallel to improve the information extraction ability of the network's shallow convolution, while the remaining layers are set as Resblock modules to fuse high-level semantic features between cross-layer convolutions. In addition, we found that the features of the first two network output layers after L2 normalization in RefineDet have not been effectively used, which reduces the detection capabilities of the latter two network output layers. In order to solve this problem, we designed the Information Reuse structure to connect the first two network output layers to the network again through the characteristics of L2 normalization. Combining the backbone RRNet we designed and the improved RefineDet network architecture, we finally got the new object detection model CXR-RefineDet, which is shown in Figure 2.

It can be seen from Figure 2 that CXR-RefineDet introduces three modules of ARM, ODM, and TCB of RefineDet network to improve the detection performance of the network. The role of the ODM module is to further accurately determine the position of the anchor frame and predict the category information of the anchor frame. The TCB module exists between the ARM and ODM modules; it integrates the context information to a greater extent by transferring the characteristics of the ARM modules in different output layers to the corresponding ODM module to improve the detection capability of the ODM module. The new backbone RRnet is designed by integrating the structure of Resblock and RepVGG block to solve the problem of poor information extraction ability of the original backbone. The structure of RRNet and Information Reuse will be further discussed in next sections.

Many natural image processing methods in computer vision have strongly relied on ImageNet pretrained deep CNN models [31] so far. These models have performed well in a large number of object categories and provide a good baseline for further model fine-tuning. In the field of object detection, the backbone usually uses a pretrained model for

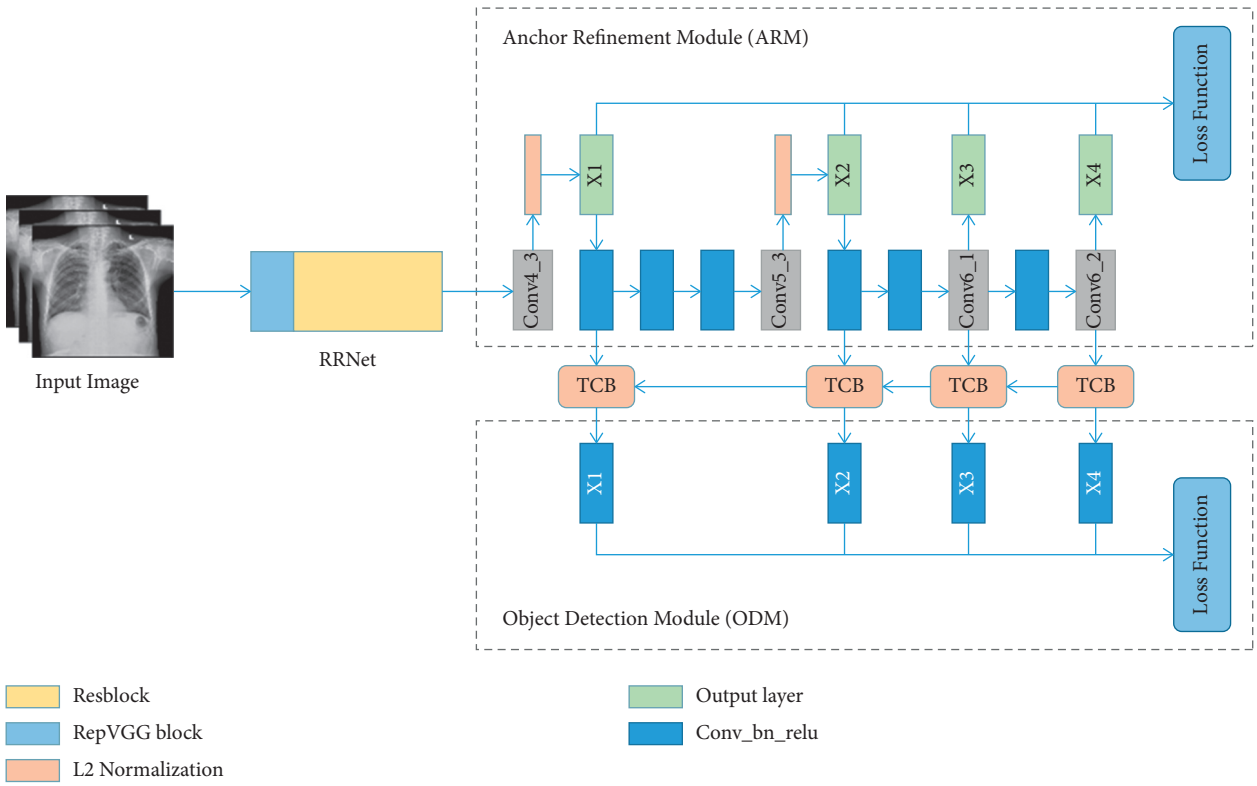


Figure 2: Architecture structure of CXR-RefineDet.

migration learning, which can accelerate the convergence speed of the network while improving the detection accuracy of the network. However, the use of pretraining models limits the flexibility of the network structure. Existing pretraining models are based on specific network structures such as ResNet-50 and ResNet-101, and their computing resource consumption often cannot meet the requirements of edge computing systems. In addition, pretraining models may not be suitable for the field of medical image diagnosis since the medical images are quite different from traditional RGB images. Using ImageNet pretraining models may cause domain mismatch problems, while training the network from scratch can avoid these problems. In summary, we follow the settings in ScratchDet [32] and introduce the BatchNorm layer in the RefineDet network, while using a larger learning rate for training.

2.3. New Design Backbone. To solve the problems of gradient dispersion and gradient explosion in deep neural network training, the residual structure proposed of ResNet introduced the jump connections in stacked convolution modules, as shown in Figure 3(b). The introduction of residual connections can fuse the feature information between different network layers, which is very beneficial for the detection of medical images. As we know, the neural network extraction of image information is a process from shallow to deep [33]. The size of the feature map will decrease as the number of network layers deepens, and the deeper the number of network layers, the more delicate the semantic information contained in the feature map. Therefore, the residual connection can further merge the fine-grained features between different convolutional layers. For example, the position of the heart and the lung lobes are included in the lung X-ray image, so the disease of cardiac hypertrophy includes the texture of the lung lobes, and the residual connection can correlate this part of the characteristic information well.

In addition, the introduction of the residual connection will not increase the model parameters since the residual connection is only the summation operation of the feature information between different network layers. As shown in Figure 3(a), unlike the cross-layer information fusion in ResNet, the multibranch topology of RepVGG is paralleled with 1×1 convolution (additional BN layer) and BN layer on both sides of 3×3 convolution (additional BN layer). Through summing the feature information extracted by different convolution modules, the information extraction capability of a single convolution module can be improved.

In our opinion, Resblock belongs to cross-layer information fusion since it merges fine-grained features between different layers through jump connections between layers. The RepVGG block connects multiple convolutional layers in parallel on a single convolution module to enrich the feature information contained in the single-layer convolution, which belongs to information fusion within the layer. Therefore, we combined Resblock and RepVGG block to design a new backbone RRNet, which can improve the information extraction capabilities of the backbone in single-layer convolution and cross-layer convolution. The network structure is shown in Figure 4(b).

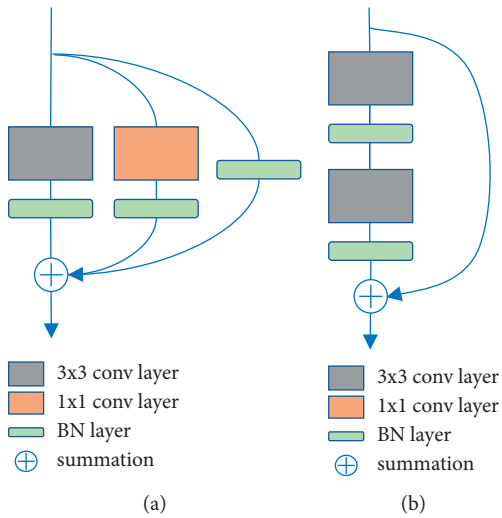


Figure 3: The architecture of different blocks: (a) RepVGG block and (b) Resblock.

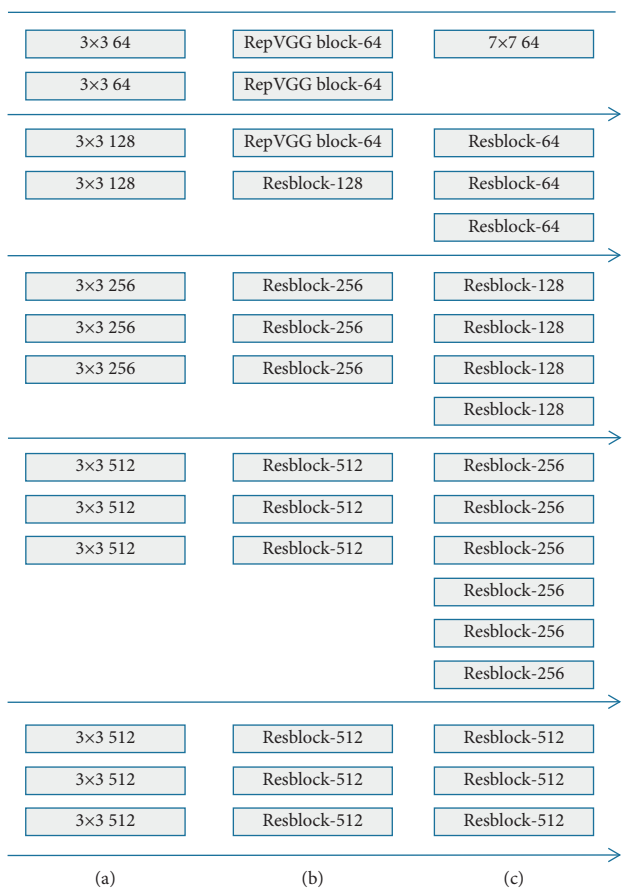


Figure 4: Comparison of VGG-16, ResNet-34, and the backbone RRNet we designed. (a) VGG-16. (b) RRNet. (c) ResNet-34.

The network shallow convolution is responsible for extracting low-level semantic features such as grayscale and texture, and the information richness of this part of the low-level semantic features directly determines the effectiveness of the high-level semantic features of the subsequent convolutional layer. With this in mind, we set the first three layers of the network as RepVGG block and use the feature of multiple modules in parallel to improve the information extraction ability of the network's shallow convolution, while the remaining layers are set as Resblock modules to fuse high-level semantic features between cross-layer convolutions. The backbone RRNet we designed has fewer network layers and higher accuracy than ResNet-34.

2.4. The Architecture of Information Reuse. We follow the settings in RefineDet and select four convolutional layers (Conv4_3, Conv5_3, Conv6_1, and Conv6_2) of different sizes in the backbone as the output of the network, as shown in Figure 5. Conv4_3 and Conv5_3 pass through L2 normalization as the first two outputs of the network, and Conv6_1 and Conv6_2 add two additional convolutional layers at the end of the VGG-16 network as the last two outputs of the network.

Although L2 normalization is added to con4_3 and conv5_3 to scale the feature norms, the scaled features are not further utilized in subsequent networks, so only the object detection effects of the first two networks output feature maps x_1 and x_2 have been improved. The structural information in the image is continuous in space and time. The shallow features extracted by the network are the representation of the deep features of the image. Using the shallow feature information extracted at the beginning of the network will help to improve the effectiveness of the subsequent extraction of deep features. Therefore, we connect con4_3 and conv5_3 to the network again after the scaling feature of L2 normalization, as shown in Figure 5. After this operation, the subsequent neural network can obtain the scaling characteristics of the output layers of the first two networks, which can improve the detection capability of the overall network.

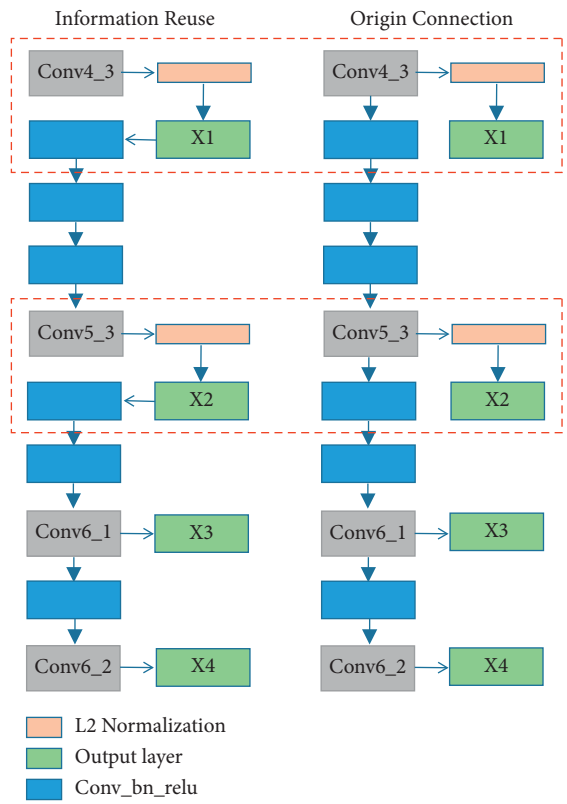


Figure 5: The connection method of the original RefineDet network output layer and the improved Information Reuse.

3. Results and Discussion

3.1. VinDr-CXR Dataset. In order to learn to annotate lesions, a large-scale and diverse lesion image dataset is required. Existing lesion datasets are usually either too small or insufficiently diverse. Fortunately, the recently released dataset VinDr-CXR greatly alleviates this limitation. VinDr-CXR is a chest radiograph dataset released by Vingroup Big Data Institute (VinBigdata) that has the most local labels and the richest number of categories so far. VinBigdata collected more than 100,000 chest radiographs from two major hospitals in Vietnam and invited 17 professional radiologists to manually label 18,000 images, of which 22 types of lesions are local tags and 6 types of special lesions are global label. In addition, VinDr-CXR dataset is divided into 15,000 training sets and 3000 test sets. The images in the training set are independently annotated by 3 doctors, and the images in the test set are jointly annotated by 5 doctors. Since 8 of the 22 categories containing local location information have a small number of images, we merge these 8 smaller categories into other lesions, and then our task is finally defined as an object detection problem for 14 types of lesions, which includes (1) aortic enlargement, (2) atelectasis, (3) calcification, (4) cardiomegaly, (5) consolidation, (6) interstitial lung disease (ILD), (7) infiltration, (8) lung opacity, (9) nodule/mass, (10) other lesions, (11) pleural effusion, (12) pleural thickening, (13) pneumothorax, and (14) pulmonary fibrosis. The distribution of each lesion category is shown in Figure 6.

3.2. Training Settings. Because the imaging principle of natural image is very different from that of medical image, the application effect of the pretraining model in medical image is not as good as the model trained from scratch. Therefore, we adopt the pretraining model and the corresponding parameter settings are fine-tuned according to ScratchDet and the experimental results, whose earning rate is 0.05, using SGD with 0.0005 weight decay and 0.9 momentum. Other training strategies mostly follow RefineDet, including data augmentation, hard negative mining, scale and aspect ratios for default boxes, and loss functions. All conv-layers are initialized with the xavier uniform method. The training of other networks such as Reitinanet and Faster RCNN is based on the mmdetection framework, and the Imagenet pretraining model and default parameters are used for training.

3.3. Ablation Study. In order to verify the effectiveness of the backbone RRNet and Information Reuse structure, we conduct ablation experiments on VinDr-CXR. In addition, we conducted comparative experiments on mainstream object detection models and compared the performance of various parameters of the models to prove the superiority of the proposed model. Specifically, we resize all the images in the training set to 512×512 resolution, but the test set is not resized, and the relevant hyperparameters for model training are kept consistent to ensure a fair comparison. All models are trained and tested on the official training set and test set and submitted to the official platform for result evaluation. The detection results were evaluated using standard PASCAL VOC 2010 [34] mean Average Precision (mAP) at IoU > 0.4.

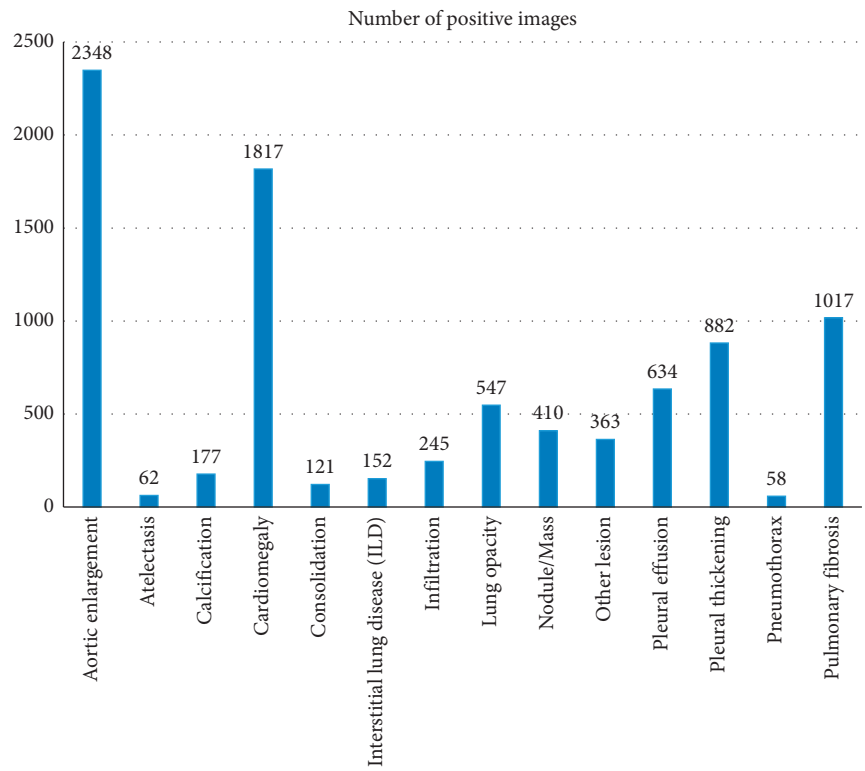


Figure 6: Distribution of various lesion categories in VinDr-CXR dataset.

Table 1: Performance comparison of different backbone networks.

Backbone	mAP	Params (M)
All-Resblock	0.1521	49.98
All-RepVGG block	0.1513	35.15
ResNet-18	0.1409	30.16
ResNet-34	0.1525	40.26
REVGG-A0	0.1374	**23.52**
REVGG-A1	0.1409	29.77
REVGG-B0	0.1434	30.23
REVGG-B1	0.1488	72.05
RRNet (ours)	**0.1572**	49.76

3.4. RRnet. We did two comparative experiments to validate the effectiveness of our new design backbone. A comparative experiment is to replace all layers of VGG-16 with Resblock and RepVGG block, and the other is to replace with two different depths of ResNet and four different versions of RepVGGnet VGG-16, as shown in Table 1. The experimental results in the first row of Table 1 show that the number of layers of the backbone RRNet we designed is less than that of ResNet-34, but the accuracy is 0.47% higher. This proves that, in our analysis in Section 2.3, the information fusion of the Resblock and RepVGG block modules in the cross-layer convolution and single-layer convolution modules can well correlate the characteristics of the lesions, thereby improving the ability of the network to extract feature information. We use Resblock and RepVGG block to replace all layers of VGG-16 and name them All-Resblock and All-RepVGG block, respectively. The two backbones obtained accuracy of 0.1521 and 0.1513, respectively, which are both higher than the ResNet-18 backbone, and the performance is equivalent to that of the ResNet-34 backbone, which verifies the effectiveness of the two modules.

In addition, it can be seen from Table 1 that the accuracy values of RepVGG-A0, A1, RepVGG-B0, and B1 are 0.1374, 0.1409, 0.1434, and 0.1488, respectively, which are lower than the accuracy value of directly replacing all layers of VGG-16 with RepVGG block. The reason is that although the RepVGG series backbones all include RepVGG block, their network tasks are designed for classification, so the difference between tasks causes the performance of the network to deteriorate. The backbone RRnet with 0.1572 mAP is higher than all versions of the backbone, which proves the effectiveness of RRnet.

3.5. Information Reuse. To demonstrate the effectiveness of the Information Reuse in the network, we use the network connection output in the original RefineDet and our

Table 2: Effectiveness of various designs.

Component	Chest-X_ray RefineDet				
VGG Pretraining model				√	
Batch normalization	√	√	√		
Information Reuse	√	√			
RRNet	√				
mAP	**0.1686**	0.1587	0.1515	0.1514	0.1173

Table 3: Detection results of different methods on VinDr-CXR test set.

Method	Backbone	Input size	mAP	Inference speed (fps) w/o reparam	Params (M)
RetinaNet	ResNet-50	512 × 512	0.1269	3.7	36.37
	ResNet-101		0.1494	3.5	55.37
Faster RCNN	ResNet-50	512 × 512	0.1597	3.7	41.19
	ResNet-101		0.1569	3.5	60.18
GA-Faster RCNN	ResNet-50	512 × 512	0.1464	3.3	41.78
Cascade RCNN	ResNet-50	512 × 512	0.1512	3.3	68.97
	ResNet-50-DCN		0.1613	3.2	69.55
	ResNet-101		0.1631	3.1	87.96
	ResNet-101-DCN		0.1560	3.0	89.24
VFNet	ResNet-50	512 × 512	0.1413	3.6	32.51
	ResNet-101		0.1505	3.4	51.51
ATSS	ResNet-50	512 × 512	0.1478	3.7	31.92
	ResNet-101		0.1538	3.5	50.91
RefineDet	VGG-16	320 × 320	0.1149	10.8	—
RefineDet	VGG-16	512 × 512	0.1173	9.9	33.51
CXR-RefineDet (ours)	RRNet	320 × 320	0.1392	9.6	—
CXR-RefineDet (ours)	RRNet	512 × 512	0.1618	6.8	49.76

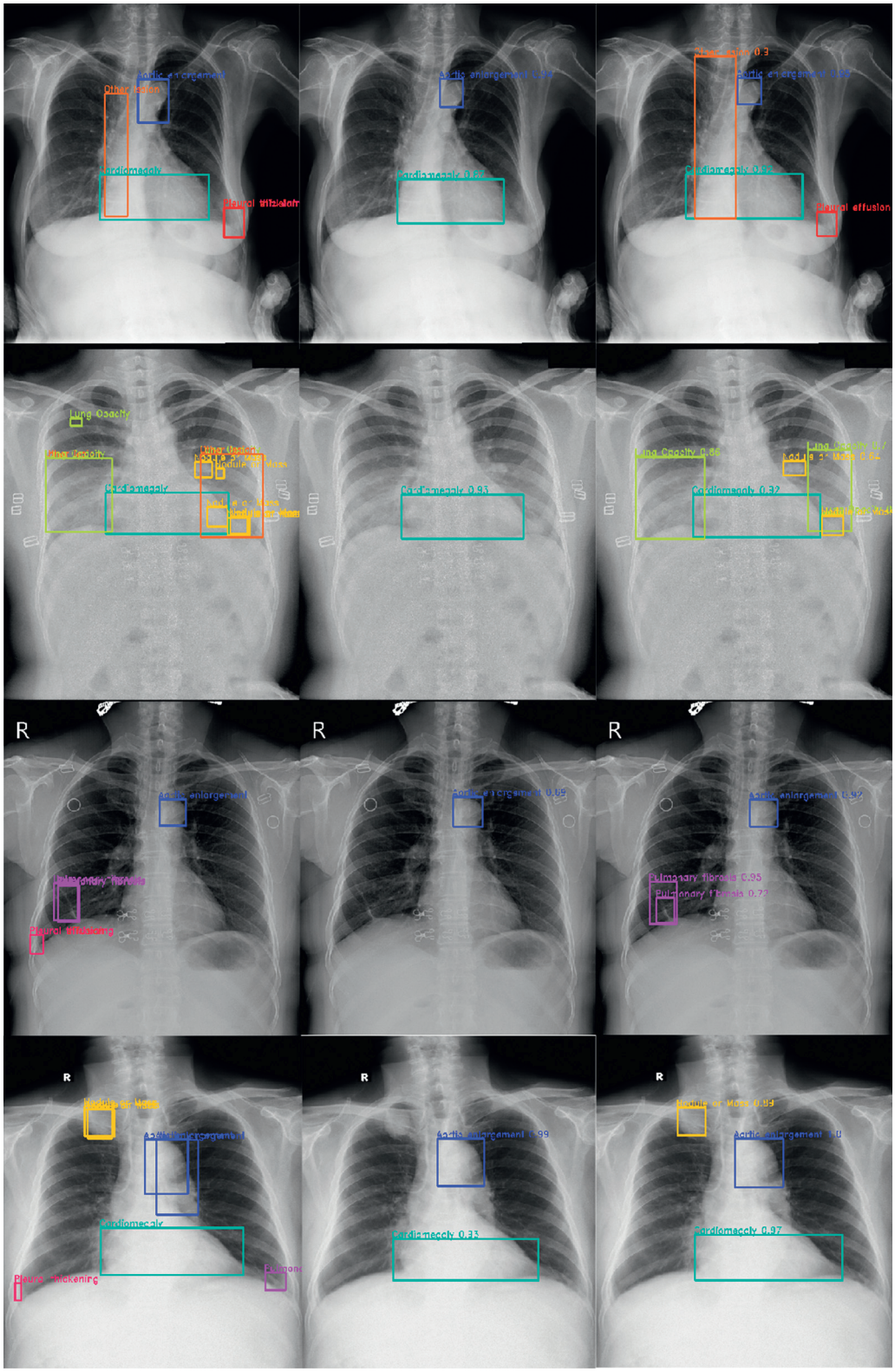
improved information reuse structure for comparative experiments, whose results are shown in the second and third columns of Table 2. After adding an improved Information Reuse structure to the basic network, the mAP is 0.1587. Compared with the basic network, adding the Information Reuse structure can bring an improvement of 0.72% mAP to the detector, which proves that reconnecting the zoom feature of the network output layer to the network through the Information Reuse structure can improve the performance of the detector.

3.6. Network Performance. In order to prove the superiority of the model, we use the three performance indicators of mAP, inference speed, and parameter quantity to conduct comparative experiments on the mainstream object detection model, and the experimental results are shown in Table 3. Compared with large-scale backbone networks such as ResNet-50 and ResNet-101, the model RefineDet with VGG-16 backbone greatly exceeds other mainstream object detectors in detection speed. The low-resolution version with 320 × 320 size of RefineDet has a detection speed of 10.8 fps, and the high-resolution version with 512 × 512 size has a detection speed of 9.9 fps. However, the low-resolution and high-resolution detection speeds of CXR-RefineDet using RRNet as the backbone are 9.6 fps and 6.8 fps, respectively.

In terms of detection accuracy, the low-resolution version of CXR-RefineDet obtained 0.1392 mAP, which surpassed RetinaNet with ResNet-50 backbone. The high-resolution version with 512 × 512 size of CXR-RefineDet obtained the highest detection accuracy of 0.1686 mAP. Compared with the multistage object detection algorithm Cascade RCNN, the model parameters of CXR-RefineDet are smaller than its backbone (ResNet-50, ResNet-101), and the accuracy is better than all its submodels. We also compare with the classic two-stage object detection algorithms Faster RCNN and GA-Faster RCNN. Although the model parameters of CXR-RefineDet with ResNet-101 backbone are slightly higher, it achieves better results in terms of detection accuracy and inference speed. In addition, we have also conducted comparative experiments with anchor-free object detection methods. VFNet [35] and ATSS [36] based on ResNet-101 backbone are slightly higher than CXR-RefineDet in model parameters, but CXR-RefineDet is better than the two methods in terms of speed and accuracy.

Based on the analysis of the above experimental results, it can be seen that the backbone RRNet and Information Reuse structure can effectively improve the detection accuracy of the network. Compared with the use of ResNet-50 and ResNet-101 as the object detector of the backbone, CXR-RefineDet not only has fewer parameters but also can reach a higher and faster level in detection accuracy and speed.

3.7. Analysis of Lesion Detection Results. The comparison diagram of detection results between RefineDet and improved CXR-RefineDet network is shown in Figure 7. It can



(1) GroundTruth-label (2) RefineDet (3) CXR-RefineDet

Figure 7: The chest X-ray test results, where (1) is the true label of the VinDr-CXR dataset, (2) and (3) are the comparison images of the test results of RefineDet and CXR-RefineDet on the VinDr-CXR test set, respectively, and the target boxes of the same color represent the same lesion category.

be seen that the detection effect of CXR-RefineDet is better than that of RefineDet network. In the detection results of the first and third rows, RefineDet cannot detect the lesions on the edge of the lung, and its detection performance is also poor for large lesions such as spine distortion. The detection result of CXR-RefineDet is similar to the truth box, and there is no missed detection or false detection. For the detection of some small targets, the detection results in the second row show that the detection rate of small targets of CXR-RefineDet is much higher than that of RefineDet. However, due to the small input resolution with 512×512, some smaller lesion areas may also be missed. For some lesions that are both small targets and extreme aspect ratios, RefineDet and CXR-RefineDet have low detection rates for these lesions, such as the red true target box in the lower left corner of the fourth row. Because the imaging principle of medical image is more complex, it does not have better discrimination than natural image. Moreover, due to the equipment, it is easy to be doped with noise in the imaging process, which brings great difficulties to the image detection. For areas with small lesions, even experienced doctors need a long time to distinguish them through naked eye observation. In addition, in order to reduce the complexity of the model, the proposed network model has fewer layers and smaller parameters, which limits the improvement of model detection ability.

4. Conclusions

In order to solve the problem of weak feature extraction capability of the RefineDet backbone network and low feature utilization of the output feature layer, a high-precision and fast detection speed lung lesion detection network CXR-RefineDet is proposed in this paper. By combining the advantages of RepVGG block and Resblock in single-layer and multilayer convolution modules, we designed an efficient backbone which was named RRNet. In view of the situation that the original network output features are not used, we introduce the Information Reuse structure to reconnect the features of the network output layer back to the network to improve the detection ability of the subsequent network. CXR-RefineDet is tested on VinDr-CXR dataset for object detection, and the detection accuracy and inference speed of CXR-RefineDet have reached 0.1686 mAP and 6.8 fps, respectively. The experimental results show that both the backbone RRNet and the structure Information Reuse can effectively improve the detection accuracy of the network. Through comparison experiments with mainstream object detection algorithms, it is found that the detection accuracy and detection speed of CXR-RefineDet are significantly better than the existing mainstream object detectors under the condition of moderate parameters. In addition, CXR-RefineDet has a good performance between accuracy and speed, which can not only effectively alleviate the problem of high computational resource consumption caused by the use of large models and large resolution in the current lung lesion detection but also provide objective conditions for the actual implementation of the computer-aided diagnosis system.

Data Availability

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that there are no conflicts of interest associated with the manuscript.

Authors' Contributions

C. Lin, Y. Zheng, X. Xiao, and J. Lin contributed equally to this work.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant no. 62072121, Natural Science Foundation of Guangdong Province (2021A1515011847), and Higher Education Reform Key Project of Hainan Province (Hnjg2018ZD-11).

References

[1] Y. Zhang, S. Wang, H. Zhao, Z. Guo, and D. Sun, "CT image classification based on convolutional neural network," *Neural Computing and Applications*, vol. 33, pp. 8191–8200, 2021.

[2] T. Matsumoto, K. Doi, A. Kano, H. Nakamura, and T. Nakanishi, "Evaluation of the potential benefit of computer-aided diagnosis (CAD) for lung cancer screenings using photofluorography: analysis of an observer study," *Nippon Igaku Hoshasen Gakkai Zasshi*, vol. 53, pp. 1195–207, 1993.

[3] H. Wang, X. Li, R. H. Jhaveri et al., "Sparse Bayesian learning based channel estimation in FBMC/OQAM industrial IoT networks," *Computer Communications*, vol. 176, pp. 40–45, 2021.

[4] Z. Wang, X. Jiang, J. Liu, K.-T. Cheng, and X. Yang, "Multi-task siamese network for retinal artery/vein separation via deep convolution along vessel," *IEEE Transactions on Medical Imaging*, vol. 39, no. 9, pp. 2904–2919, 2020.

[5] J. A. Tariq, I. A. Ben, and K. S. Abdel, "Tumor edge detection in mammography images using quantum and machine learning approaches," *Neural Computing and Applications*, vol. 33, no. 13, pp. 7773–7784, 2021.

[6] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580–587, Columbus, OH, USA, June 2014.

[7] M. Zhao, H. Wang, Y. Han et al., "SEENS: nuclei segmentation in Pap smear images with selective edge enhancement," *Future Generation Computer Systems*, vol. 114, 2020.

[8] A. A. A. Setio, F. Ciompi, G. Litjens et al., "Pulmonary nodule detection in ct images: false positive reduction using multi-view convolutional networks," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1160–1169, 2016.

[9] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Proceedings of the International Conference on Medical Image Computing and Computer-Aassisted Iintervention*, Munich, Germany, October 2015.

[10] J. Rocha, A. Cunha, and A. M. Mendonça, "Conventional filtering versus U-Net based models for pulmonary nodule segmentation in CT images," *Journal of Medical Systems*, vol. 44, no. 4, p. 81, 2020.

[11] Z. Q. Zhao, P. Zheng, S. T. Xu, and X. Wu, "Object detection with deep learning: a review," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, pp. 3212–3232, 2018.

[12] E. Kim, S. Kim, M. Seo, and S. Yoon, "XProtoNet: diagnosis in chest radiography with global and local explanations," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, USA, June 2014.

[13] R. Girshick, "Fast R-CNN," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 1440–1448, Santiago, Chile, December 2015.

[14] X. Li, L. Shen, and X. Xie, "Multi-resolution convolutional networks for chest X-Ray radiograph based lung nodule detection," *Artificial Intelligence in Medicine*, vol. 103, Article ID 101744, 2019.

[15] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 2961–2969, Venice, Italy, October 2017.

[16] K. Yan, Y. Peng, V. Sandfort, M. Bagheri, and Z. Lu, "RM summers holistic and comprehensive annotation of clinically significant findings on diverse CT images: learning from radiology reports and label ontology," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, 2019.

[17] J. Liang, G. Ye, J. Guo, Q. Huang, and S. Zhang, "Reducing false-positives in lung nodules detection using balanced datasets," *Frontiers in Public Health*, vol. 9, 2021.

[18] Y. Xiao, X. Wang, Q. Li et al., "A cascade and heterogeneous neural network for CT pulmonary nodule detection and its evaluation on both phantom and patient data," *Computerized Medical Imaging and Graphics*, vol. 90, Article ID 101889, 2021.

[19] T. Y. Lin, P. Dollar, and R. Girshick, "Feature pyramid networks for object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, July 2017.

[20] T. Y. Lin, P. Goyal, R. Girshick, K. M. He, and P. Dollar, "Focal loss for dense object detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 99, pp. 2999–3007, 2017.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 770, Las Vegas, NV, USA, June 2016.

[22] H. Jie, S. Li, S. Gang, and S. Albanie, "Squeeze-and-excitation networks," in *Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 99, Salt Lake City, UT, USA, June 2017.

[23] H. Nguyen, K. Lam, L. T. Le et al., "VinDr-CXR: an open dataset of chest X-rays with radiologist's annotations," 2012, https://arxiv.org/abs/2012.15029.

[24] S. Zhang, L. Wen, X. Bian, Z. Lei, and S. Z. Li, "Single-shot refinement neural network for object detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4203–4212, Salt Lake City, UT, USA, June 2018.

[25] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, https://arxiv.org/abs/1409.1556.

[26] G. Huang, Z. Liu, L. Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4700–4708, Honolulu, HI, USA, July 2017.

[27] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, July 2016.

[28] H. Zhang, C. Wu, Z. Zhang et al., "Resnest: splitattention networks," 2020, https://arxiv.org/abs/2004.08955.

[29] X. Li, W. Wang, X. Hu, and J. Yang, "Selective kernel networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, June 2020.

[30] X. Ding, X. Zhang, N. Ma, J. Han, G. Ding, and J. Sun, "RepVGG: making VGG-style ConvNets great again," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, USA, June 2021.

[31] J. Deng, W. Dong, R. Socher, L. Li, K. Li, and F. Li, "Imagenet: a large--scale hierarchical image database," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, Miami, FL, USA, June 2009.

[32] R. Zhu, F. Zhang, X. Wang, L. Wen, and T. Mei, "ScratchDet: training single-shot object detectors from scratch," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, June 2019.

[33] Q. Chen, M. Huang, H. Wang, and G. Xu, "A feature discretization method based on fuzzy rough sets for high-resolution remote sensing big data under linear spectral model," *IEEE Transactions on Fuzzy Systems*, vol. 2021, Article ID 3058020, 1 page, 2021.

[34] M. Everingham, S. M. A. Eslami, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes challenge: a retrospective," *International Journal of Computer Vision*, vol. 111, no. 1, pp. 98–136, 2015.

[35] H. Zhang, Y. Wang, F. Dayoub, and N. Sünderhauf, "VarifocalNet: an IoU-aware dense object detector," 2020, https://arxiv.org/abs/2004.08955.

[36] S. Zhang, C. Chi, Y. Yao, Z. Lei, and S. Z. Li, "Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, June 2020.

Hindawi
Journal of Healthcare Engineering
Volume 2022, Article ID 3774123, 12 pages
https://doi.org/10.1155/2022/3774123

Research Article

Solving an Infectious Disease Model considering Its Anatomical Variables with Stochastic Numerical Procedures

Zulqurnain Sabir,[1] **Muhammad Asif Zahoor Raja**,[2] **and Yolanda Guerrero Sánchez**[3]

[1]*Department of Mathematics and Statistics, Hazara University, Mansehra, Pakistan*
[2]*Future Technology Research Center, National Yunlin University of Science and Technology, 123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan*
[3]*Department of Anathomy and Pscicobiology, Faculty of Medicine, University of Murcia, Murcia 30100, Spain*

Correspondence should be addressed to Yolanda Guerrero Sánchez; yolanda.guerreros@um.es

Received 12 October 2021; Revised 13 December 2021; Accepted 20 December 2021; Published 7 January 2022

Academic Editor: Han Wang

The aim of the current work is to perform the numerical investigation of the infectious disease based on the nonlinear fractional order prey-predator model using the Levenberg–Marquardt backpropagation (LMB) based on the artificial neuron networks (ANNs), i.e., LMBNNs. The fractional prey-predator model is classified into three categories, the densities of the susceptible, infected prey, and predator populations. The statistics proportions for solving three different variations of the infectious disease based on the fractional prey-predator model are designated for training 80% and 10% for both validation and testing. The numerical actions are performed using the LMBNNs to solve the infectious disease based on the fractional prey-predator model, and comparison is performed using the database Adams–Bashforth–Moulton approach. The infectious disease based on the fractional prey-predator model is solved using the LMBNNs to reduce the mean square error (M.S.E). In order to validate the exactness, capability, consistency, and competence of the proposed LMBNNs, the numerical procedures using the correlation, M.S.E, regression, and error histograms are drawn.

1. Introduction

Infectious diseases occur when some viruses, fungi, germs, and parasites enter into the human body. These forms are diffused through infection from one to another human, contaminated food, animals, or contact to any of the ecological factors that are polluted with any type of these bodies. Every infection-based disease has its own symptoms, types, and severity. A few common symptoms of these infections in the body include pain, flu, cough, and fever [1, 2]. Some of the infections have minor symptoms that do not need any cure or treatment. Alternatively, there are various serious deathly cases that may disturb the population equilibrium of numerous classes in the atmosphere. For the last few years, mathematical systems are used to predict the species evolution. It initiated from the Lotka and Volterra systems [3, 4], where their expediency in evading several worst situations for many species as death was evidenced. Currently, researchers apply this tool to reveal the consequence of a certain policy used by some of the governments to handle few species that can be measured as a significant device to preserve each kind.

The ecological systems are more complex for any type of infection, which can affect the growth of few classes, as a personification. In this study, a predator-prey collaboration is considered. This contagion may distress the predator strength and the hunting competence that takes few of the predators in the position of death. In the previous studies, numerous investigations have examined the predator-prey interactions in the occurrence of transmittable viruses; see, for instance, [5, 6] where predator-prey models are treated with an analytical approach, while in [7–9], these models have a numerical analysis. As an alternative, there are various approaches, which reflect the predators to

accomplish an effective hunt. The hunting collaboration is one of the operating policies of the predator, where numerous predators work to hunt some prey. This scheme is valuable to reduce the rate of hunting, and a number of hunters behave in this technique. The high rate of efficiency is seen in hunting lions, wild dogs, and hyenas. The mathematical system of this precise predator behavior was modeled and presented for the first time in [10], in which a simple model was applied to describe a cooperation. So far, there have been a few studies of the predator-prey interaction behavior presented in these references: in [11], the hunting effect is considered, in [12], the cooperation effect is considered, in [13], hunting and cooperation effects are considerd at the same time, and finally, in [14], hunting and cooperation effects are considered jointly with the Alle effect. Therefore, it has investigated the effects of a transmittable virus in the predator-prey communication along with the occurrence of the collaboration of predator hunting.

A three-species system is considered an infection in a prey population, which is categorized into two classes, the infected and susceptible prey. It is found that the derivative forms of time fractional have widespread applications to describe various forms of actual conditions that are recognized by the memory outcomes of a dynamical form. The memory rate is known as the derivative order, and the function of memory is called the fractional-order derivative. The time-fractional derivative is implemented to model the phenomena of various real-world problems [15, 16]. The nonlinear fractional-order prey-predator model has three classes, mathematically written as follows [17]:

$$\begin{cases} D^{\alpha}S(\tau) = r(I(\tau)+S(\tau)) - (aP(\tau)+\lambda)P(\tau)S(\tau) - \delta S(\tau)I(\tau) - \mu S(\tau), & S(0)=k_1, \\ D^{\alpha}I(\tau) = \delta I(\tau)S(\tau) - (\mathrm{aP}(\tau)+\lambda)P(\tau)I(\tau) - \mu I(\tau), & I(0)=k_2, \\ D^{\alpha}P(\tau) = e(\mathrm{aP}(\tau)+\lambda)P(\tau)(I(\tau)+S(\tau)) - \mathrm{mP}(\tau), & P(0)=k_3, \end{cases} \tag{1}$$

where $S(\tau)$ and $I(\tau)$ are the densities of the susceptible and infected prey, while $P(\tau)$ shows the densities of predator populations. The term e signifies the rate of conversion of prey into predator biomass (susceptible or infected). The parameter r represents the prey population's reproduction number, and it is supposed that this infection does not convey vertically. It can also be defined as mother-child, where the predator is not diseased by this virus after a direct interaction with the diseased prey. The rate of transmission of the prey population, i.e., infection rate, is represented by δ. The functional form $(\mathrm{aP}(\tau)+\lambda)P(\tau)S(\tau)$ and $(\mathrm{aP}(\tau)+\lambda)P(\tau)I(\tau)$ are the hunting cooperation values [10]. The death rate of the population of the prey is μ, which represents the natural mortality rate of the predator's population. The initial conditions are k_1, k_2, and k_3, respectively.

The aim of the current work is to perform the numerical investigation of the infectious disease based on the fractional prey-predator model using the Levenberg–Marquardt backpropagation (LMB) based on the artificial neuron networks (ANNs), i.e., LMBNNs. The LMBNNs are applied on three different variants of authentication, testing, training, and sample information. The statistics proportions for solving three different variations of the infectious disease based on the fractional prey-predator model are designated for training 80% and 10% for both validation and testing. The numerical results are performed using the LMBNNs to solve the infectious disease based on the fractional prey-predator model, and comparison is performed using the database Adams–Bashforth–Moulton approach. Recently, the stochastic computing solvers are applied based on the heuristic and swarming techniques in frequently reported articles of utmost significance [18–21]. However, the infectious disease spread systems governed with the fractional prey-predator model have never implemented to study its solution dynamics by using the competency of LMBNNs' computing paradigm. A few novel features and contribution of the current investigations are provided in brief as follows:

(i) The design of stochastic computing solvers LMBNNs is presented for the first time to solve the infectious disease spread systems governed with the fractional prey-predator model

(ii) The designed procedures of LMBNNs have been implemented effectively to study the behavior of different scenarios of the fractional prey-predator model, and comparative studies are found in decent resemblance with the state-of-the-art Adams–Bashforth–Moulton numerical approach for solving fractional differential equations

(iii) The convergence performances on iterative updated of MSE, negligible absolute error (AE) from standard outcomes, correlation/regression index, and error histograms (EHs) further authenticate the efficacy of the designed LMBNN computing platform for fractional prey-predator models

The paper is organized as follows: Section 2 shows the methodology based on LMBNNs. Section 3 presents the numerical outcomes through LMBNNs to solve the fractional-order nonlinear prey-predator model. The final comments are reported in Section 4.

2. Methodology: LMBNNs

In this section, the proposed methodology of LMBNNs is presented for the infectious disease based on the fractional-order nonlinear prey-predator system. The methodology is categorized in two steps. The necessary trials of the stochastic-based LMBNNs are provided, and the execution

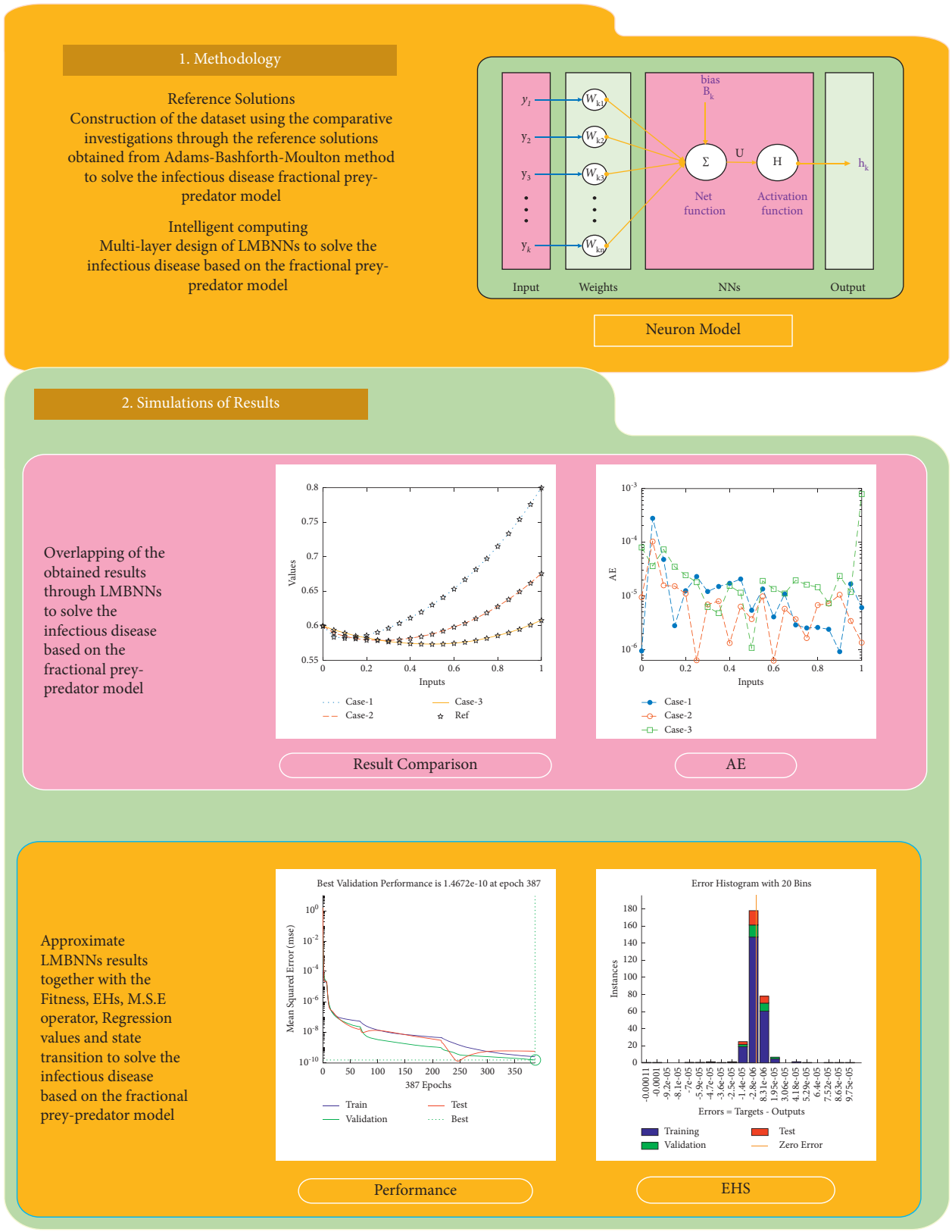


Figure 1: Workflow illustration of the designed LMBNNs to solve the infectious disease based on the nonlinear fractional prey-predator model.

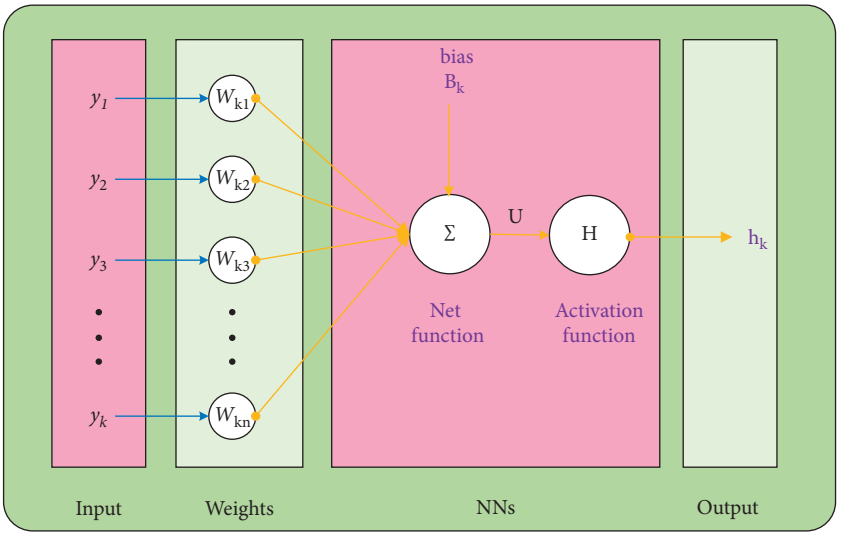


Figure 2: Proposed framework using a single neuron.

process of the stochastic computing scheme is given to solve the infectious disease based on the nonlinear fractional-order prey-predator model.

A suitable optimization procedure-based proposed LMBNN is plotted in Figure 1 together with the outcomes and analysis of the results, while for a single neuron, the designed procedure is given in Figure 2. The stochastic computing-based procedures are executed using the "nftool" (MATLAB build-in command in the neural networks toolbox). The dataset for the fractional-order nonlinear prey-predator system is designated for training 80% and 10% for both validation and testing in LMBNN operations.

3. Numerical Procedures of the Fractional-Order Nonlinear Prey-Predator System

The current section shows the numerical procedures of the infectious disease based on the fractional-order nonlinear prey-predator system by applying the proposed computing stochastic LMBNNs. The literature parameter forms to solve the infectious disease based on the nonlinear fractional-order prey-predator model are $r = 1.5$, $\lambda = 0.5$, $a = 0.5$, $\delta = 0.5$, $\mu = 0.5$, $e = 0.5$, $m = 0.5$, $k_1 = 0.2$, $k_2 = 0.7$, and $k_3 = 0.6$. Three cases using the variations of fractional-order derivative, i.e., $\alpha = 0.5$, 0.7, and 0.9, are provided to solve the infectious disease based on the fractional-order nonlinear prey-predator system. The inclusive results have been performed for each category of the fractional-order nonlinear prey-predator system which are in between [0, 1] with 0.01 step size. Ten numbers of neurons throughout this numerical study have been taken, and the data are designated for training 80% and 10% for both validation and testing. The achieved numerical values using the LMBNNs to solve the infectious disease based on the fractional-order nonlinear prey-predator system are drawn in Figure 3. The representations based on the LMBNNs to solve the infectious disease based on the fractional prey-predator system are given in Figures 4–8. The M.S.E measures and state transitions (STs) to solve the infectious disease based on the fractional prey-predator model are plotted in Figure 4. The M.S.E based on the training, states of best curves, authentication, and testing is drawn in Figures 4(a)–4(c)), whereas the best ST values to solve the fractional prey-predator model are derived in Figures 4(d)–4(f)) at epochs 325, 387, and 127, respectively. The obtained performances exist around 1.3399×10^{-10}, 1.4672×10^{-10}, and 5.2351×10^{-10}, respectively. The gradient performances of the LMBNNs to solve the infectious disease based on the fractional prey-predator model are found around 8.0331×10^{-06}, 9.9161×10^{-08}, and 9.8524×10^{-08}, respectively. These calculated performances plotted in the figures show the accuracy, convergence, and precision of the proposed stochastic procedures of the LMBNNs to solve the infectious disease based on the fractional prey-predator model. The plots of the fitting curves to solve the fractional prey-predator model are given in Figures 5(a)–5(c)), which show the comparative analysis of the obtained outcomes through LMBNNs. Figure 5(d)–5(f)) show the values of the EHs that exists around 6.63×10^{-07}, 8.31×10^{-06}, and 3.19×10^{-05} for the 1st, 2nd, and 3rd case, respectively. The values of the regression are drawn in Figures 6–8 to solve the infectious disease based on the fractional prey-predator model. These illustrations of the correlations indicate regression soundings found around 1 that authenticates the perfect model. The testing, verification, and training plots designate the exactness of the LMBNNs to solve the infectious disease based on the fractional prey-predator model. In addition, the convergence performances through M.S.E based on the epochs, training, complexity, testing, backpropagation performances, and verification are provided in Table 1 to solve the infectious disease based on the nonlinear fractional prey-predator model.

The comparative performances and the AE values are illustrated in Figures 9 and 10 for the fractional prey-predator model. The outcomes for each category of the fractional prey-predator system presented using the stochastic LMBNNs are given in Figures 9(a)–9(c)). The matching of the obtained and reference solutions for each

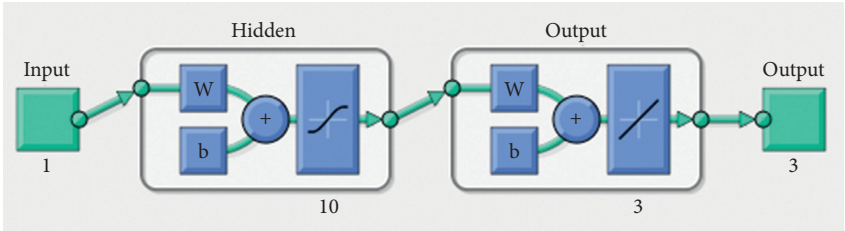


Figure 3: Proposed LMBNN to solve the infectious disease based on the nonlinear fractional prey-predator model.

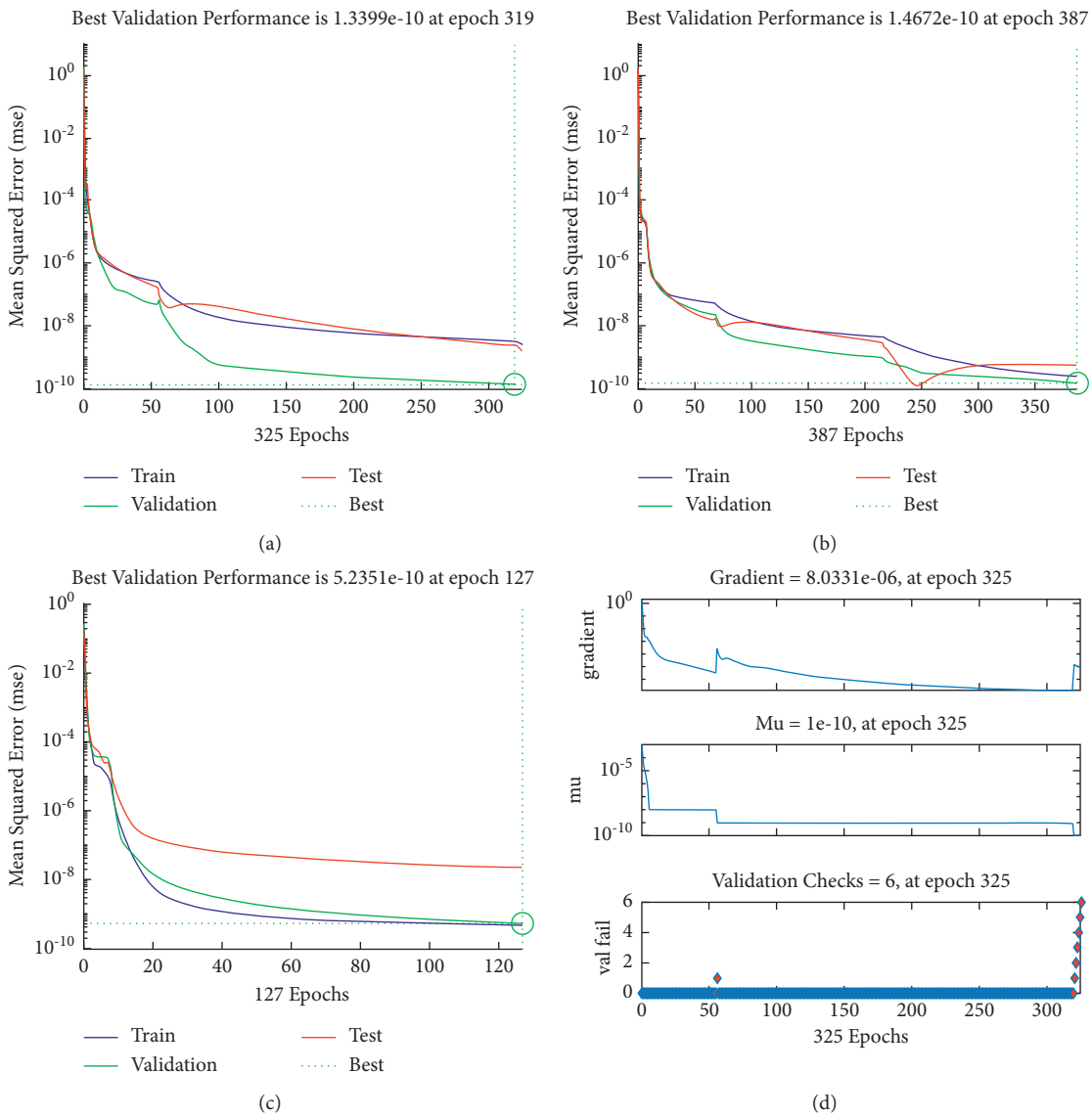


Figure 4: Continued.

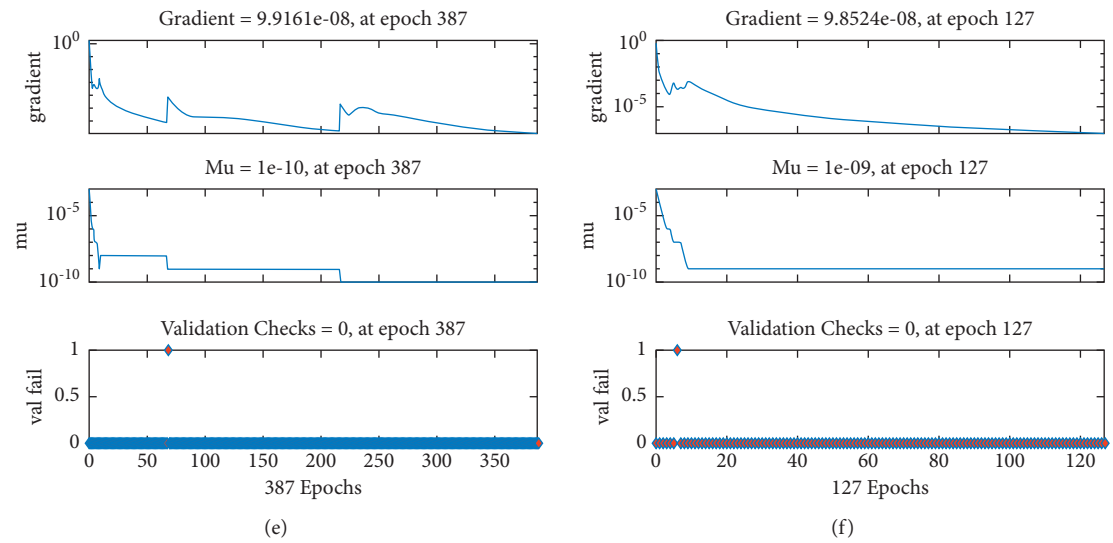


Figure 4: M.S.E measures (a–c) and ST performances (d–f) to solve the infectious disease based on the fractional prey-predator model. (a) Case 1: M.S.E. (b) Case 2: M.S.E. (c) Case 3: M.S.E. (d) Case I: EHs. (e) Case 2: EHs. (f) Case 3: EHs.

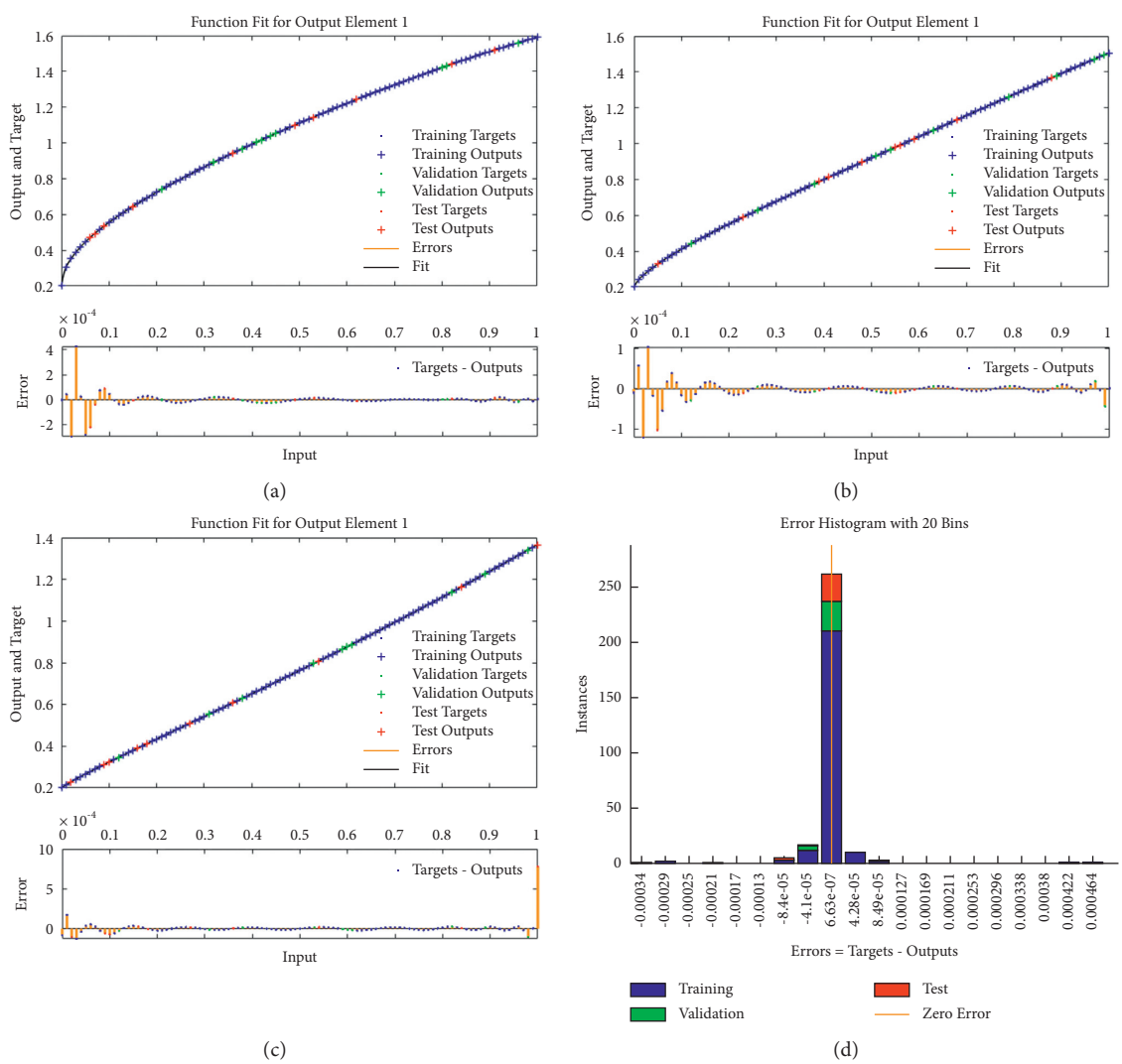


Figure 5: Continued.

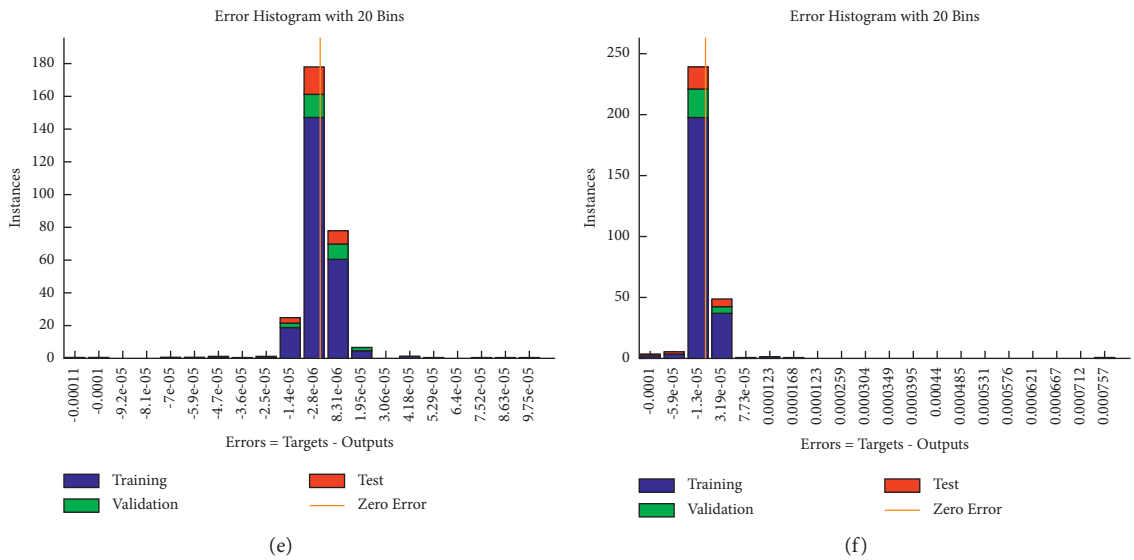


Figure 5: Result assessments (a–c) and EHs (d–f) to solve the infectious disease based on the fractional prey-predator model. (a) Case 1: result assessments. (b) Case 2: result assessments, (c) Case 3: result assessments. (d) Case I: EHs. (e) Case 2: EHs. (f) Case 3: EHs.

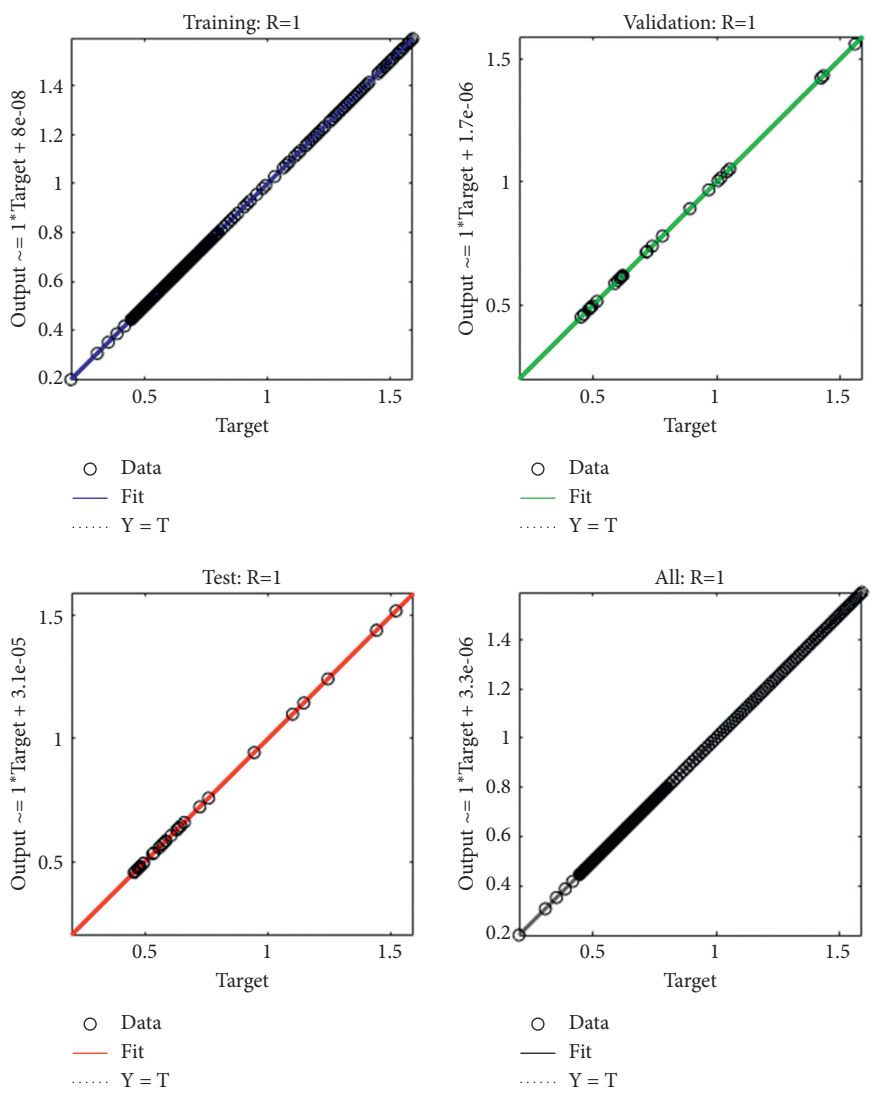


Figure 6: Regression measures of the fractional-order nonlinear prey-predator model of case 1.

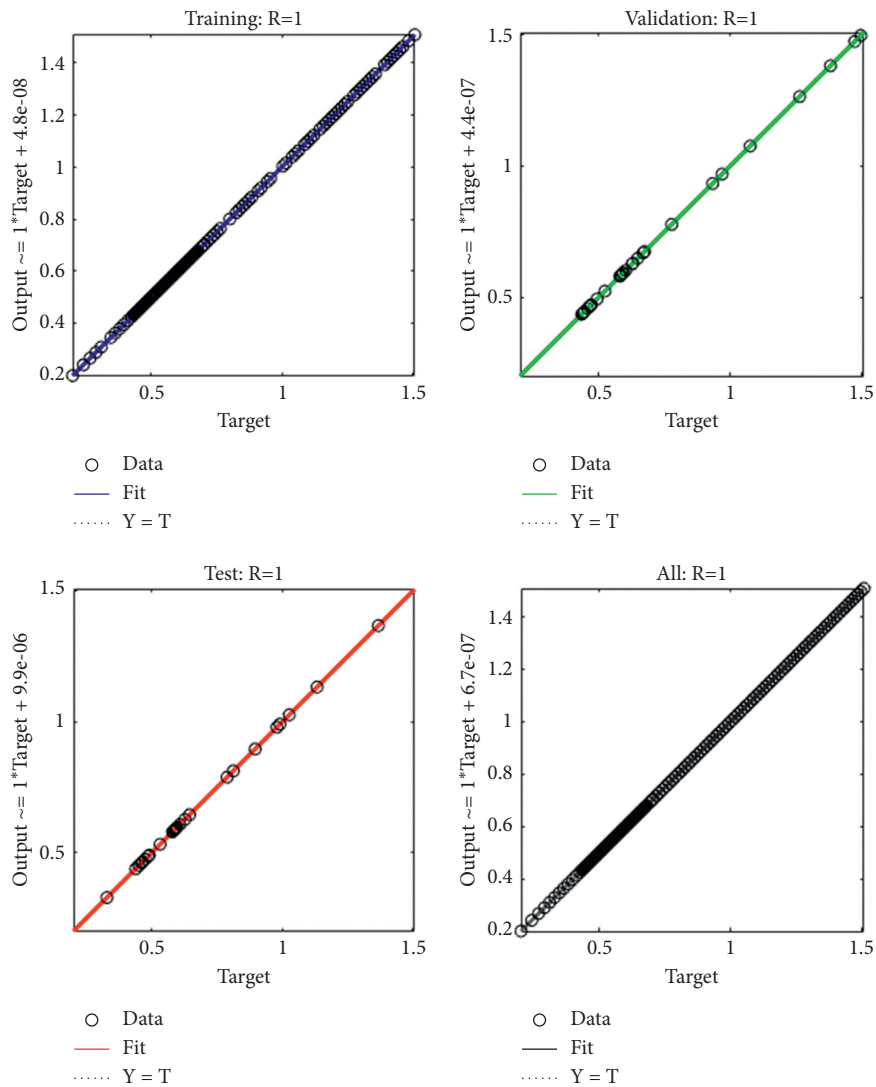


Figure 7: Regression measures of the fractional-order nonlinear prey-predator model of case 2.

category of the infectious disease is perceived based on the fractional-order nonlinear prey-predator model. These outcomes matching represent the accurateness of the stochastic LMBNNs for each category of the infectious disease based on the fractional order nonlinear prey-predator system, The AE measures for each category of the infectious disease based on the fractional-order nonlinear prey-predator system are plotted in Figures 10(a)–10(c)). The AE for $S(\tau)$ based on the fractional-order nonlinear prey-predator system is calculated around 10^{-04} to 10^{-06} for case 1 and 3, while the AE is calculated 10^{-05} to 10^{-06} for case 3. The AE for $I(\tau)$ based on the fractional-order nonlinear prey-predator model is calculated around 10^{-04} to 10^{-07} for case 1, 10^{-05} to 10^{-09} for case 2, and 10^{-05} to 10^{-07} for case 3. The AE

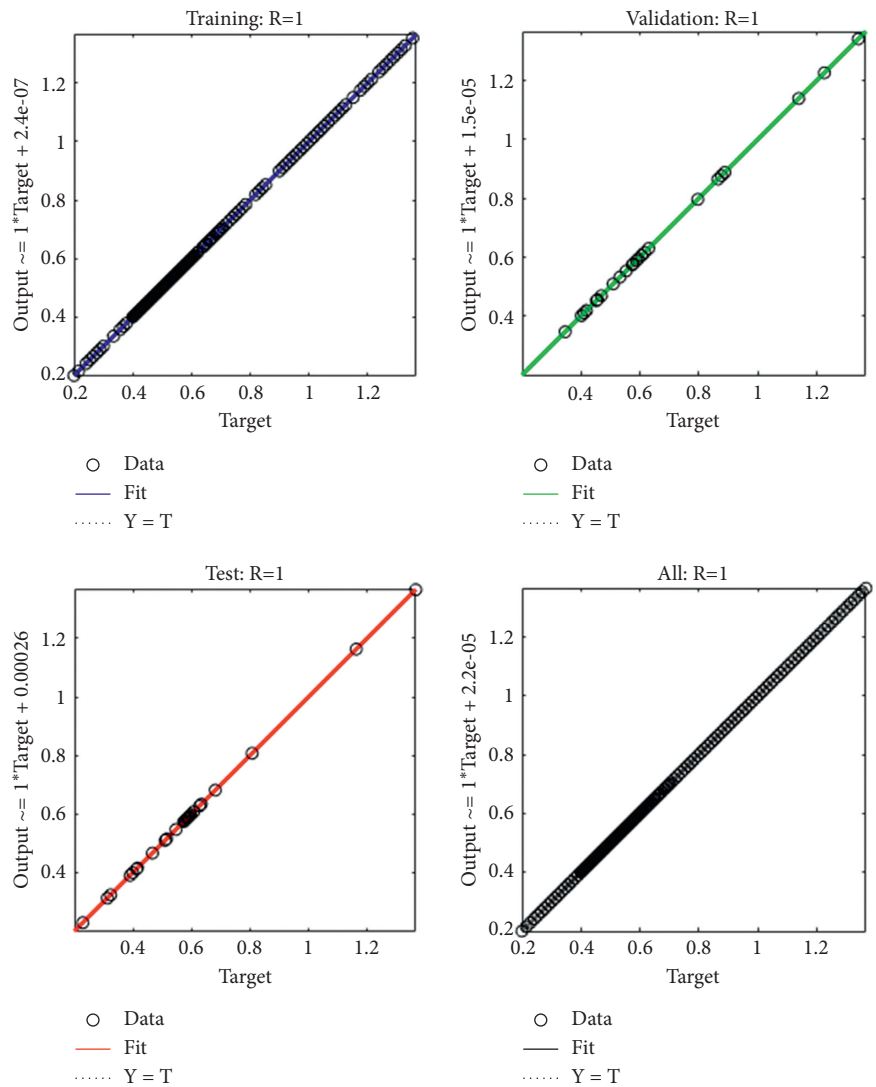


Figure 8: Regression measures of the fractional-order nonlinear prey-predator model of case 3.

Table 1: Statistical measures to solve the infectious disease based on the nonlinear fractional prey-predator model.

Case	M.S.E for samples in			Performance	Gradient	Mu	Epoch	Time
	Training	Validation	Testing					
1	3.21×10^{-09}	1.33×10^{-10}	2.29×10^{-09}	2.63×10^{-09}	8.03×10^{-06}	1.00×10^{-10}	325	5
2	2.43×10^{-10}	1.46×10^{-10}	5.44×10^{-10}	2.43×10^{-10}	9.92×10^{-08}	1.00×10^{-10}	387	5
3	4.63×10^{-10}	5.23×10^{-10}	2.19×10^{-08}	4.63×10^{-10}	9.85×10^{-08}	1.00×10^{-09}	127	2

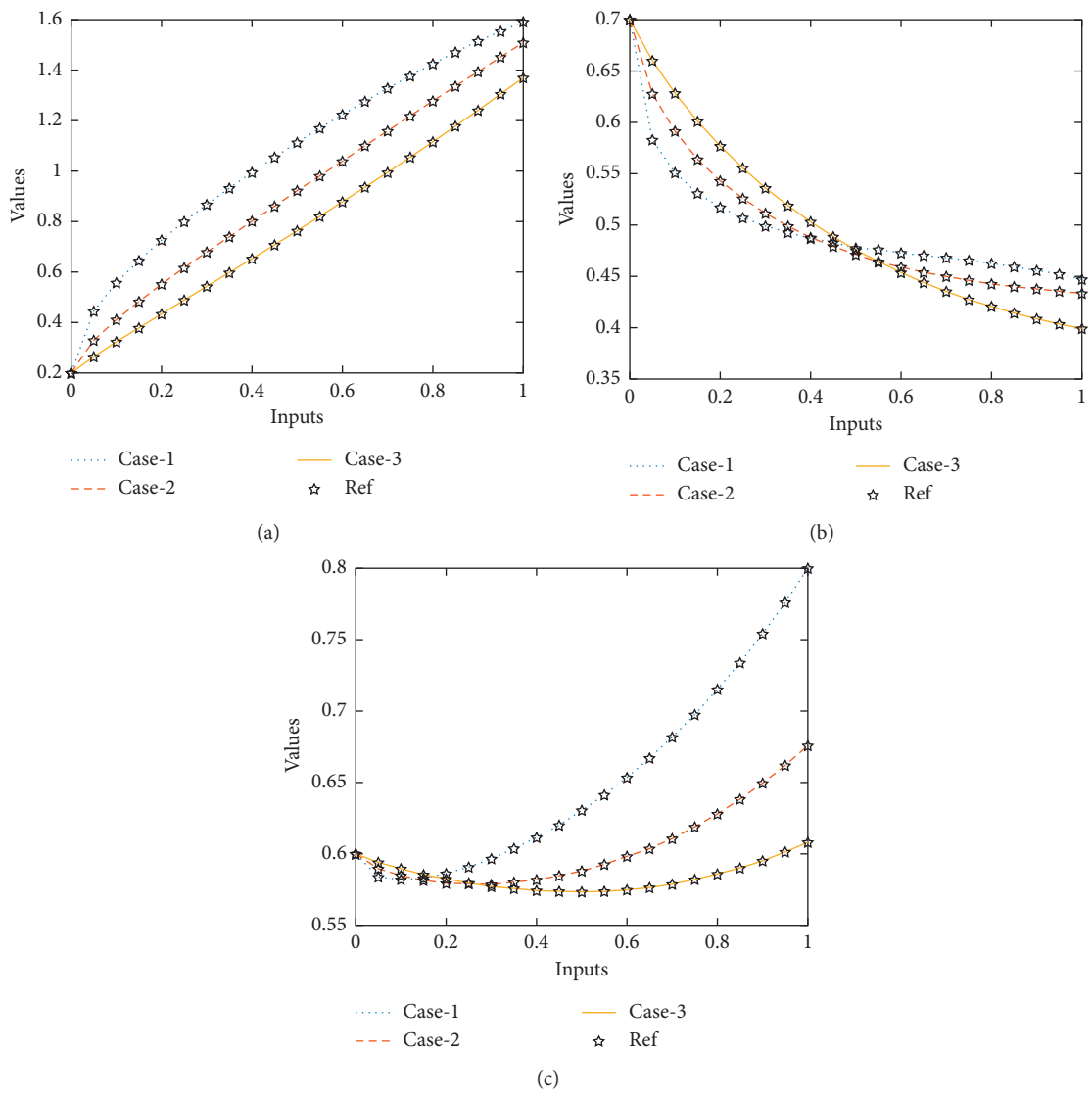


Figure 9: Comparison illustrations to solve the infectious disease based on the nonlinear fractional prey-predator model. (a) Results for $S(\tau)$. (b) Results for $I(\tau)$. (c) Results for $P(\tau)$.

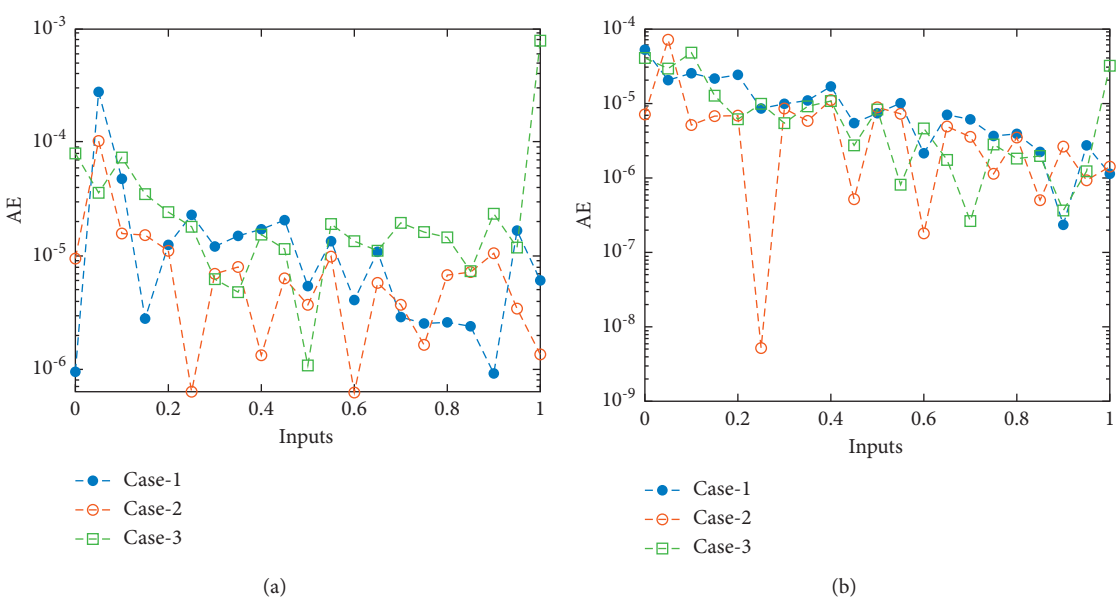
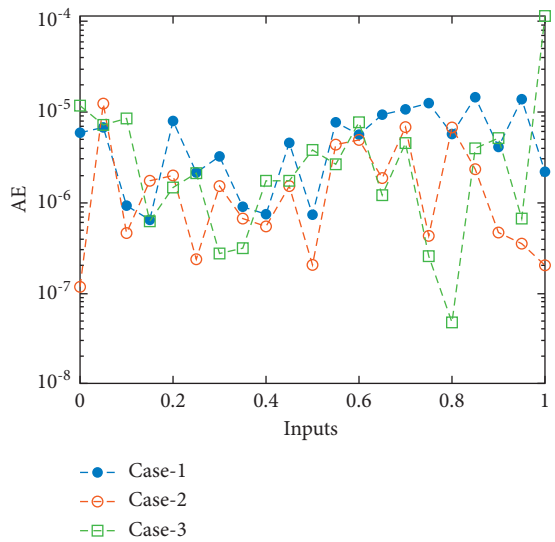


Figure 10: Continued.



(c)

Figure 10: AE values to solve the infectious disease based on the nonlinear fractional prey-predator model. (a) AE for $S(\tau)$. (b) AE for $I(\tau)$. (c) AE for $P(\tau)$.

for $P(\tau)$ based on the fractional-order nonlinear prey-predator model is calculated around 10^{-05} to 10^{-08} for case each case of the nonlinear system.

4. Conclusions

In these investigations, the solutions of an infectious virus based on the nonlinear fractional prey-predator system are numerically presented by using the stochastic procedures based on the Levenberg–Marquardt backpropagation along with artificial neural networks. These stochastic-based procedures LMBNNs are provided to solve three cases by taking different values of the fractional order. The numerical solutions have been performed using the sample data, testing, training, and authentication. The data proportions to solve the nonlinear fractional prey-predator model are designated for training 80% and 10% for both validation and testing. The numerical results of the infectious disease based on the nonlinear fractional prey-predator model are achieved using the LMBNNs, and comparison is performed using the database Adams–Bashforth–Moulton approach. The solutions of the fractional-order nonlinear model are obtained through the LMBNNs in order to reduce the M.S.E. To indorse the exactness, capability, dependability, and competence of the proposed LMBNNs, the numerical procedures are provided using the M.S.E, correlation, EHS, and regression. The matching of the results designates the precision of the designed scheme, and the values of the AE in good ranges for each case of the infectious disease based on the nonlinear fractional prey-predator model show the effectiveness of the scheme.

In future, the procedures based on the LMBNNs are applied to get the outcomes of the fractional-order systems and Lonngren-wave systems [22–26]. Additionally, one may exploit the Bayesian regularization method-based neural networks for solving different scenarios of the fractional prey-predator model for better outcomes in terms of accuracy and efficiency.

Data Availability

No data were used to support this study.

Conflicts of Interest

All authors declare no potential conflicts of interest.

Authors' Contributions

All authors have worked in an equal way to develop this work. Prof. Sabir and *Raja* have developed the mathematical part, and Prof. Guerrero has led the medical modeling.

Acknowledgments

This paper was partially supported by Ministerio de Ciencia, Innovación y Universidades (No.PGC2018-097198-B-I00) and Fundación Séneca de la Región de Murcia (No.20783/PI/18).

References

[1] Y. G. Sánchez, Z. U. L. Q. U. R. N. A. I. N. Sabir, and J. U. A. N. L. G. Guirao, "Design of a nonlinear SITR fractal model based on the dynamics of a novel coronavirus (COVID-19)," *Fractals*, vol. 28, no. 08, Article ID 2040026, 2020.

[2] Z. Sabir, M. U. H. A. M. M. A. D. Umar, M. U. H. A. M. M. A. D. A. S. I. F. Z. A. H. O. O. R. Raja, and D. U. M. I. T. R. U. Baleanu, "Applications of Gudermannian neural network for solving the SITR fractal system," *Fractals*, 2021.

[3] A. J. Lotka-Volterra, "Elements of physical biology," *Nature*, vol. 116, 1925.

[4] V. Volterra, *SUI TENTATIVI DI APPLICAZIONE DELLE MATEMATICHE ALLE SCIENZE BIOLOGICHE E SOCIALI DISCORSO LETTO IL 4 NOVEMBRE 1901 alla inaugurazione dell'anno scolastico nella R. Università di Roma DAL PROF. VITO VOLTERRA*, pp. 436–458, Giornale degli economisti, 1901.

[5] J. Chattopadhyay and O. Arino, "A predator-prey model with disease in the prey," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 36, no. 6, pp. 747–766, 1999.

[6] K. P. Hadeler and H. I. Freedman, "Predator-prey populations with parasitic infection," *Journal of Mathematical Biology*, vol. 27, no. 6, pp. 609–631, 1989.

[7] M. Umar, Z. Sabir, and M. A. Z. Raja, "Intelligent computing for numerical treatment of nonlinear prey-predator models," *Applied Soft Computing*, vol. 80, pp. 506–524, 2019.

[8] L. Han, Z. Ma, and H. W. Hethcote, "Four predator prey models with infectious diseases," *Mathematical and Computer Modelling*, vol. 34, no. 7-8, pp. 849–858, 2001.

[9] X. Zhou, J. Cui, X. Shi, and X. Song, "A modified Leslie–Gower predator–prey model with prey infection," *Journal of Applied Mathematics and Computing*, vol. 33, no. 1, pp. 471–487, 2010.

[10] J. Duarte, C. Januário, N. Martins, and J. Sardanyés, "Chaos and crises in a model for cooperative hunting: a symbolic dynamics approach," *Chaos*, vol. 19, no. 4, Article ID 043102, 2009.

[11] S. Pal, N. Pal, S. Samanta, and J. Chattopadhyay, "Effect of hunting cooperation and fear in a predator-prey model," *Ecological Complexity*, vol. 39, Article ID 100770, 2019.

[12] K. Ryu and W. Ko, "Asymptotic behavior of positive solutions to a predator–prey elliptic system with strong hunting cooperation in predators," *Physica A: Statistical Mechanics and Its Applications*, vol. 531, Article ID 121726, 2019.

[13] T. Singh, fnm au, R. Dubey, and V. Narayan Mishra, "Spatial dynamics of predator-prey system with hunting cooperation in predators and type I functional response," *AIMS Mathematics*, vol. 5, no. 1, pp. 673–684, 2020.

[14] D. Song, Y. Song, and C. Li, "Stability and turing patterns in a predator–prey model with hunting cooperation and allee effect in prey population," *International Journal of Bifurcation and Chaos*, vol. 30, no. 09, Article ID 2050137, 2020.

[15] B. Ghanbari and S. Djilali, "Mathematical and numerical analysis of a three-species predator-prey model with herd behavior and time fractional-order derivative," *Mathematical Methods in the Applied Sciences*, vol. 43, no. 4, pp. 1736–1752, 2020.

[16] B. Ghanbari and S. Djilali, "Mathematical analysis of a fractional-order predator-prey model with prey social behavior and infection developed in predator population," *Chaos, Solitons & Fractals*, vol. 138, Article ID 109960, 2020.

[17] S. Djilali and B. Ghanbari, "The influence of an infectious disease on a prey-predator model equipped with a fractional-order derivative," *Advances in Difference Equations*, vol. 2021, no. 1, pp. 1–16, 2021.

[18] K. Nisar, "Evolutionary integrated heuristic with gudermannian neural networks for second kind of lane-emden nonlinear singular models," *Applied Sciences*, vol. 11, no. 11, p. 4725, 2021.

[19] M. K. Ammar, M. R. Amin, and M. H. M. Hassan, "Calculation of line of site periods between two artificial satellites under the action air drag," *Applied Mathematics and Nonlinear Sciences*, vol. 3, no. 2, pp. 339–352, 2018.

[20] K. Nisar, Z. Sabir, M. A. Z. Raja, A. A. Ibrahim, F. Erdogan, and M. Reazul Haque, "Design of morlet wavelet neural network for solving a class of singular pantograph nonlinear differential models," IEEE Access, 2021.

[21] M. K. Ammar, M. R. Amin, and M. H. M. Hassan, "Visibility intervals between two artificial satellites under the action of Earth oblateness," *Applied Mathematics and Nonlinear Sciences*, vol. 3, no. 2, pp. 353–374, 2018.

[22] E. Ilhan and İ O. Kıymaz, "A generalization of truncated M-fractional derivative and applications to fractional differential equations," *Applied Mathematics and Nonlinear Sciences*, vol. 5, no. 1, pp. 171–188, 2020.

[23] H. M. Baskonus, H. Bulut, and T. A. Sulaiman, "New complex hyperbolic structures to the lonngren-wave equation by using sine-gordon expansion method," *Applied Mathematics and Nonlinear Sciences*, vol. 4, no. 1, pp. 141–150, 2019.

[24] Z. Sabir, A. Ayub, J. L. G. Guirao, S. Bhatti, and S. Z. H. Shah, "The effects of activation energy and thermophoretic diffusion of nanoparticles on steady micropolar fluid along with Brownian motion," *Advances in Materials Science and Engineering*, vol. 2020, Article ID 2010568, 12 pages, 2020.

[25] K. Vajravelu, S. Sreenadh, and R. Saravana, "Influence of velocity slip and temperature jump conditions on the peristaltic flow of a Jeffrey fluid in contact with a Newtonian fluid," *Applied Mathematics and Nonlinear Sciences*, vol. 2, no. 2, pp. 429–442, 2017.

[26] M. S. M. Selvi and L. Rajendran, "Application of modified wavelet and homotopy perturbation methods to nonlinear oscillation problems," *Applied Mathematics and Nonlinear Sciences*, vol. 4, no. 2, pp. 351–364, 2019.

Hindawi
Journal of Healthcare Engineering
Volume 2021, Article ID 7633381, 12 pages
https://doi.org/10.1155/2021/7633381

Research Article

Stroke Disease Detection and Prediction Using Robust Learning Approaches

Tahia Tazin,[1] Md Nur Alam,[1] Nahian Nakiba Dola,[1] Mohammad Sajibul Bari,[1] Sami Bourouis,[2] and Mohammad Monirujjaman Khan[1]

[1]*Department of Electrical and Computer Engineering, North South University, Bashundhara, Dhaka 1229, Bangladesh*
[2]*Department of Information Technology, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia*

Correspondence should be addressed to Mohammad Monirujjaman Khan; monirujjaman.khan@northsouth.edu

Received 7 October 2021; Revised 4 November 2021; Accepted 9 November 2021; Published 26 November 2021

Academic Editor: Han Wang

Stroke is a medical disorder in which the blood arteries in the brain are ruptured, causing damage to the brain. When the supply of blood and other nutrients to the brain is interrupted, symptoms might develop. According to the World Health Organization (WHO), stroke is the greatest cause of death and disability globally. Early recognition of the various warning signs of a stroke can help reduce the severity of the stroke. Different machine learning (ML) models have been developed to predict the likelihood of a stroke occurring in the brain. This research uses a range of physiological parameters and machine learning algorithms, such as Logistic Regression (LR), Decision Tree (DT) Classification, Random Forest (RF) Classification, and Voting Classifier, to train four different models for reliable prediction. Random Forest was the best performing algorithm for this task with an accuracy of approximately 96 percent. The dataset used in the development of the method was the open-access Stroke Prediction dataset. The accuracy percentage of the models used in this investigation is significantly higher than that of previous studies, indicating that the models used in this investigation are more reliable. Numerous model comparisons have established their robustness, and the scheme can be deduced from the study analysis.

1. Introduction

Stroke occurs when the blood flow to various areas of the brain is disrupted or diminished, resulting in the cells in those areas of the brain not receiving the nutrients and oxygen they require and dying. A stroke is a medical emergency that requires urgent medical attention. Early detection and appropriate management are required to prevent further damage to the affected area of the brain and other complications in other parts of the body. The World Health Organization (WHO) estimates that fifteen million people worldwide suffer from strokes each year, with one person dying every four to five minutes in the affected population. Stroke is the sixth leading cause of mortality in the United States according to the Centers for Disease Control and Prevention (CDC) [1]. Stroke is a non-communicable disease that kills approximately 11% of the population. In the United States, approximately 795,000 people suffer from the disabling effects of strokes on a regular basis [2]. It is India's fourth leading cause of death. Strokes are classified as ischemic or hemorrhagic. In a chemical stroke, clots obstruct the drainage; in a hemorrhagic stroke, a weak blood vessel bursts and bleeds into the brain. Stroke may be avoided by leading a healthy and balanced lifestyle that includes abstaining from unhealthy behaviors, such as smoking and drinking, keeping a healthy body mass index (BMI) and an average glucose level, and maintaining an excellent heart and kidney function. Stroke prediction is essential and must be treated promptly to avoid irreversible damage or death. With the development of technology in the medical sector, it is now possible to anticipate the onset of a stroke by utilizing ML techniques. The algorithms included in ML are beneficial as they allow for accurate prediction and proper analysis. The majority of

previous stroke-related research has focused on, among other things, the prediction of heart attacks. Brain stroke has been the subject of very few studies. The main motivation of this paper is to demonstrate how ML may be used to forecast the onset of a brain stroke. The most important aspect of the methods employed and the findings achieved is that among the four distinct classification algorithms tested, Random Forest fared the best, achieving a higher accuracy metric in comparison to the others. One downside of the model is that it is trained on textual data rather than real time brain images. The implementation of four ML classification methods is shown in this paper.

Numerous academics have previously utilized machine learning to forecast strokes. Govindarajan et al. [3] used text mining and a machine learning classifier to classify stroke disorders in 507 individuals. They tested a variety of machine learning methods for training purposes, including Artificial Neural Network (ANN), and they found that the SGD algorithm provided the greatest value, 95 percent. Amini et al. [4, 5] performed research to predict a stroke occurrence. They classified 50 risk variables for stroke, diabetes, cardiovascular disease, smoking, hyperlipidemia, and alcohol consumption in 807 healthy and unhealthy individuals. They used two of the most accurate methods: the c4.5 decision tree algorithm (95 percent accuracy) and the K-nearest neighbor algorithm (94 percent accuracy). Cheng et al. [6] presented a study on estimating the prognosis of an ischemic stroke. In their study, they used 82 ischemic stroke patient data sets, two ANN models, and the accuracy values of 79 and 95 percent. Cheon et al. [7–9] conducted research to determine the predictability of a stroke patient death. They identified the stroke incidence using 15,099 individuals in their research. They detected strokes using a deep neural network method. The authors utilized PCA to extract information from the medical records and predict strokes. They have 83 percent area under the curve (AUC). Singh et al. [10] conducted research using artificial intelligence to predict strokes. They employed a new technique for predicting stroke in their research using the cardiovascular health study (CHS) dataset. Additionally, they used the decision tree method to do a feature extraction followed by a principal component analysis. In this case, the model was built using a neural network classification method, and it achieved 97 percent accuracy.

Chin et al. [11] conducted research to determine the accuracy of an automated early ischemic stroke detection. The major objective of their research was to create a method for automating primary ischemic stroke using Convolutional Neural Network (CNN). They amassed 256 pictures for the purpose of training and testing the CNN model. They utilized the data lengthening technique to increase the gathered picture in their system's image preparation. Their CNN technique achieved a 90 percent accuracy rate. Sung et al. [12] conducted research to establish a stroke severity index. They gathered data on 3577 patients who had an acute ischemic stroke. They utilized a variety of data mining methods, including linear regression, to create their predictive models. Their ability to predict outperformed the k-nearest neighbor method (95% confidence interval). Monteiro et al. [13] used machine learning to predict the functional prognosis of an ischemic stroke. They tested this method on a patient who died three months after admission. They obtained an AUC value of greater than 90. Kansadub et al. [14] conducted research to determine the risk of stroke. The authors of the research analyzed the data to predict strokes using Naive Bayes, decision trees, and neural networks. They assessed their pointer's accuracy and AUC in their research. They categorized all of these algorithms as decision trees, with naive Bayes providing the most accurate results. Adam et al. [15] conducted research to determine the classification of an ischemic stroke. They categorized ischemic strokes using two models: the k-nearest neighbor method and the decision tree technique. In their study, the decision tree method was found to be more useful by medical experts when used to categorize strokes.

The majority of studies had an accuracy rate of around 90%, which was considered to be quite good. However, the novelty of our research is that we used several well-known machine learning methods to get the best result. Random forest (RF), decision tree (DT), voting classifier (VC), and logistic regression (LR) were the most successful algorithms, with 96, 94, 91, and 87 percent F1-scores, respectively. The accuracy percent of the models used in this research is much greater than the accuracy percent of the models used in previous investigations, suggesting that the models used in this investigation are more trustworthy. They have been shown to be resilient in many model comparisons, and the scheme may be generated from the results of the study's analysis.

As mentioned earlier, the major contribution of this research is that we have used different machine learning models on a publicly available dataset. In the previous work, most of the researchers used a significant model to predict the stroke disease. However, we used four different models, and also, we compared the results with the previous work. All the results and comparisons are briefly discussed in the following section. The rest of this article is set out as follows: the experimental methodology and procedures are described in Section 2; the result analysis is provided in Section 3; and conclusions have been discussed in Section 4.

2. Procedure and Experimental Methodology

This section includes a description of the dataset, a block diagram, a flow diagram, and evaluation matrices, as well as the process and methodology used in the study.

2.1. Proposed System. The data has become available for model construction once it has been processed. A preprocessed dataset and machine learning techniques are needed for the model construction. LR, DT classification, RF classification, and voting classifier are some of the methods used. After creating four alternative models, the accuracy measures, namely accuracy score, precision score, recall score, and F1 score are used to compare them. The designed system's block diagram is shown in Figure 1.

All the components of the block diagram have been discussed in the following subsections.

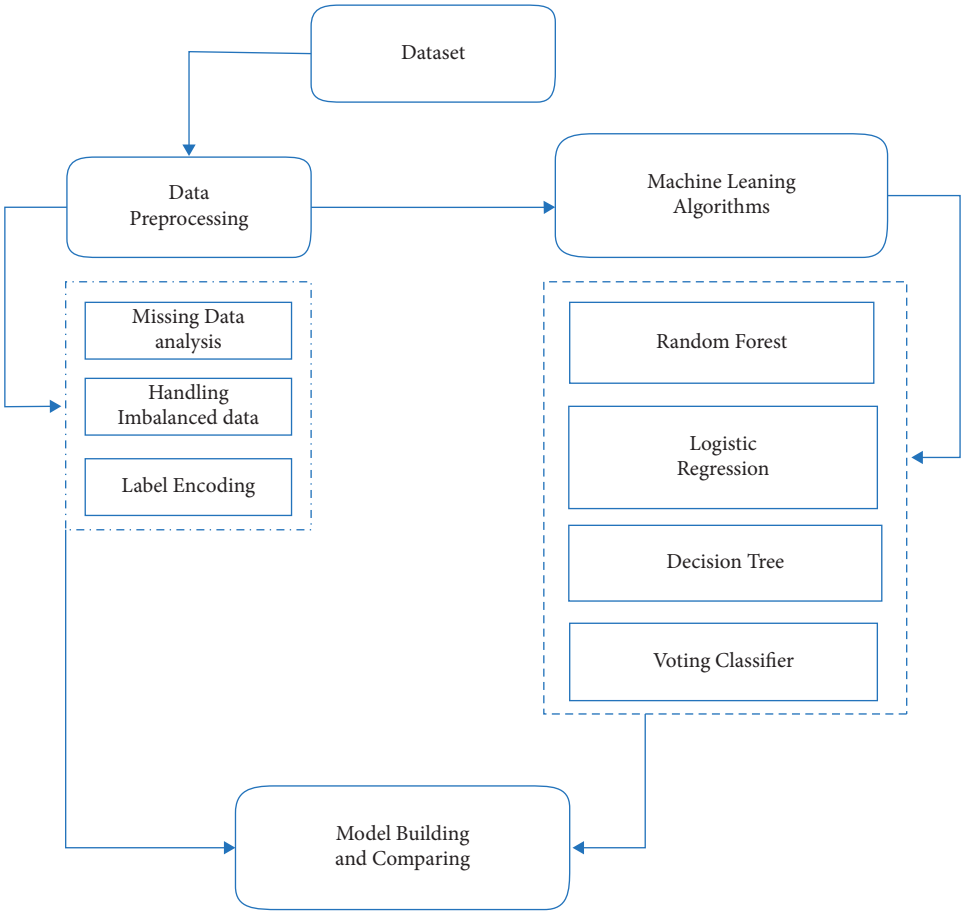


Figure 1: Proposed system's block diagram.

2.2. Dataset. The stroke prediction dataset [16] was used to perform the study. There were 5110 rows and 12 columns in this dataset. The value of the output column stroke is either 1 or 0. The number 0 indicates that no stroke risk was identified, while the value 1 indicates that a stroke risk was detected. The probability of 0 in the output column (stroke) exceeds the possibility of 1 in the same column in this dataset. 249 rows alone in the stroke column have the value 1, whereas 4861 rows have the value 0. To improve accuracy, data preprocessing is used to balance the data. Figure 2 shows the total number of stroke and nonstroke records in the output column before preprocessing.

From Figure 2, it is clear that this dataset is an imbalanced dataset. The SMOTE technique has been used to balance this dataset.

2.3. Preprocessing. Before building a model, data preprocessing is required to remove unwanted noise and outliers from the dataset that could lead the model to depart from its intended training. This stage addresses everything that prevents the model from functioning more efficiently. Following the collection of the relevant dataset, the data must be cleaned and prepared for model development. As stated before, the dataset used has twelve characteristics. To begin with, the column id is omitted since its presence has no bearing on model construction. The dataset is then inspected for null values and filled if any are detected. The null values in the column BMI are filled using the data column's mean in this case.

Label encoding converts the dataset's string literals to integer values that the computer can comprehend. As the computer is frequently trained on numbers, the strings must be converted to integers. The gathered dataset has five columns of the data type string. All strings are encoded during label encoding, and the whole dataset is transformed into a collection of numbers. The dataset used for stroke prediction is very imbalanced. The dataset has a total of 5110 rows, with 249 rows indicating the possibility of a stroke and 4861 rows confirming the lack of a stroke. While using such data to train a machine-level model may result in accuracy, other accuracy measures such as precision and recall are inadequate. If such an unbalanced data is not dealt with properly, the findings will be inaccurate, and the forecast will be ineffective. As a result, to obtain an efficient model, this unbalanced data must be dealt with first. The SMOTE technique was employed for this purpose. Figure 3 depicts the dataset's balance output column.

The next stage is to construct the model after finishing data preparation and managing the imbalanced dataset. To improve the accuracy and efficiency of this job, the data is divided into training and testing data with a ratio of 80

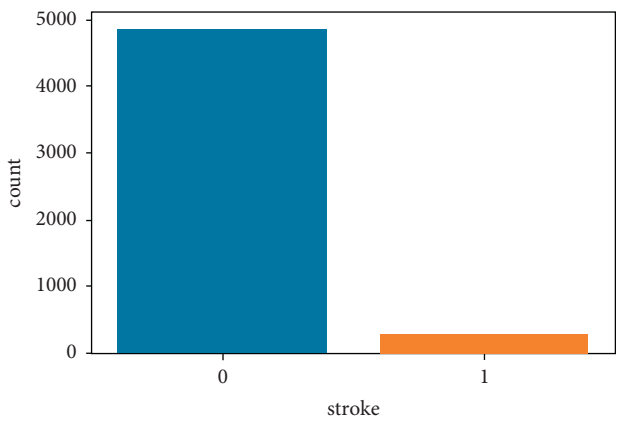


Figure 2: Total number of stroke and normal data.

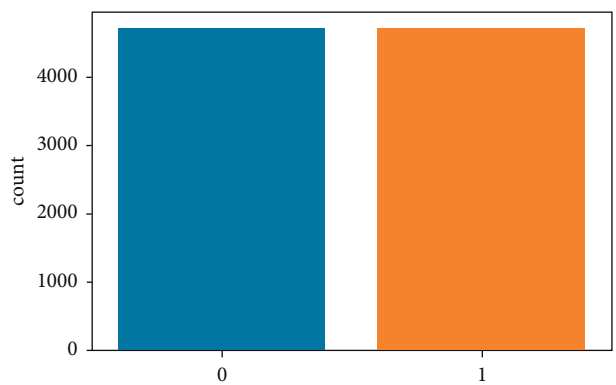


Figure 3: Output column after preprocessing.

percent training data and 20 percent testing data. After splitting, the model is trained using a variety of classification methods. Random forest, decision tree classification method, voting classifier, and logistic regression are the classification algorithms utilized in this study.

2.4. Proposed Algorithms. The most common disease identified in the medical field is stroke, which is on the rise year after year. Using the publicly accessible stroke prediction dataset, the study measured four commonly used machine learning methods for predicting brain stroke recurrence, which are as follows:

(i) Random forest
(ii) Decision tree
(iii) Voting classifier
(iv) Logistic regression

2.4.1. Random Forest. The classification algorithm chosen was RF classification [17]. RFs are composed of numerous independent decision trees that were trained individually on a random sample of data. These trees are created during training, and the decision trees' outputs are collected. A process termed voting is used to determine the final forecast made by this algorithm. Each DT in this method must vote for one of the two output classes (in this case, stroke or no stroke). The final prediction is determined by the RF method, which chooses the class with the most votes. A block diagram of random forest classification is shown in Figure 4.

The flexibility of the random forest is one of its most alluring features. It may be utilized for relapse detection and grouping tasks, and the overall weighting given to information characteristics is readily apparent. Additionally, it is a beneficial approach since the default hyperparameters it employs often give unambiguous expectations. Understanding the hyperparameters is critical since there are relatively few of them, to begin with. Overfitting is a well-known problem in machine learning, although it occurs seldom with the arbitrary random forest classifier. If there are sufficient trees in the forest, the classifier will not overfit the model.

2.4.2. Decision Tree. Both regression and classification concerns are addressed using classification with DT [18]. Furthermore, as the input variables already have a related output variable, this methodology is a supervised learning model. It resembles a tree. The data is constantly segmented according to a specific parameter in this method. The decision node and the leaf node are the two parts of a decision tree. At the former node, the data is divided, and the latter is the node that produces the result. The DT classifier's basic structure is depicted in Figure 5.

The DT is easy to comprehend since it replicates the phases that a person goes through while making a real world decision. It may be very beneficial in resolving issues with decision-making. Consider all potential solutions to an issue. Cleaning data is not required as much as it is with other methods.

2.4.3. Voting Classifier. A voting classifier is a type of classification model that trains on an ensemble of multiple models and predicts an output (class) based on the class that has the greatest chance of being selected as the output [19]. It is used to predict the outcome of a vote. The flowchart for the voting classifier model is shown in Figure 6.

Voting summarizes the methodology we will use to compare various training models. There are two methods of voting, which are as follows:

(i) Soft voting: In this phase, the predicted probability gradients for each model are added and averaged. The category with the highest value is deemed the winner, and its contents are the output. While this seems to be a fair and rational strategy, it is only recommended if the individual categories are calibrated correctly. This is similar to computing the weighted average of a set of numbers, except that each of the various models contributes proportionally to the final output vector.
(ii) Hard voting: This phase combines the categorization outputs of all the various models and specifies the

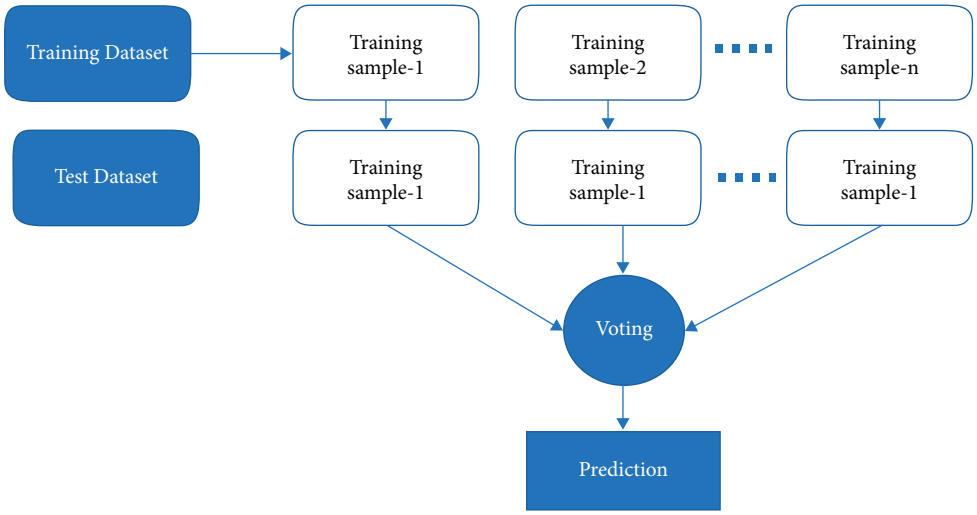


Figure 4: Block diagram of Random Forest Classifier.

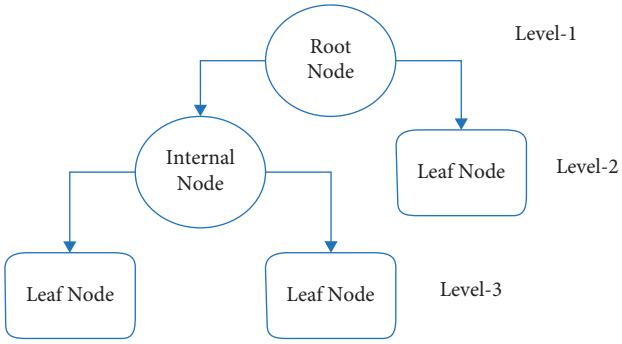


Figure 5: Basic structure of a decision tree classifier.

final output value as the mode value of the resultant output. Because of the fact that the particular probability values associated with each model are disregarded, this approach is analogous to computing the arithmetic mean of a collection of numbers. The output alone of each model is considered.

2.4.4. Logistic Regression. The flowchart for the logistic regression model is shown in Figure 7. In the supervised learning approach, LR is one of the most commonly used ML algorithms [20]. It is a forecasting method that uses a collection of independent factors to predict a categorical dependent variable.

Utilizing logistic regression, the output of a categorical dependent variable is predicted. As a result, the output must be discrete or categorical in nature. It may be yes or no, 0 or 1, true or false, etc., but probability values between 0 and 1 are given. Logistic regression and linear regression are used in very similar ways. The classification problems are addressed with LR, and the regression problems are addressed using linear regression. Instead of a regression line, we use an S-shaped logistic function that predicts the two maximum values (0 or 1).

2.5. Evaluation Matrix. Figure 8 depicts the confusion matrix or evaluation matrix. The confusion matrix is a tool for evaluating the performance of machine learning classification algorithms. The confusion matrix has been used to test the efficiency of all models created. The confusion matrix illustrates how often our models forecast correctly and how often they estimate incorrectly. False positives and false negatives have been allocated to badly predicted values, whereas true positives and true negatives were assigned to properly anticipated values. The model's accuracy, precision-recall trade-off, and AUC were utilized to assess its performance after grouping all predicted values in the matrix.

3. Result Analysis

The models' capacities, model forecasts, investigation, and eventual outcomes are examined in this part.

3.1. Data Visualization. A histogram depicts a recurrence dispersion with infinite classes. It is a region outline made of square shapes with bases at class boundary spans and regions proportionate to the comparing classes' frequencies. As the base fills in the spaces between the class borders, the square

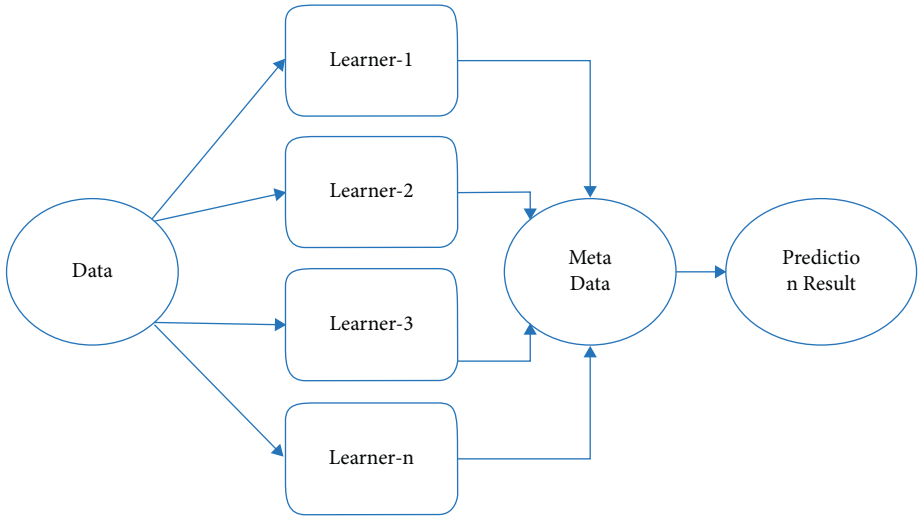


Figure 6: Flowchart of a voting classifier.

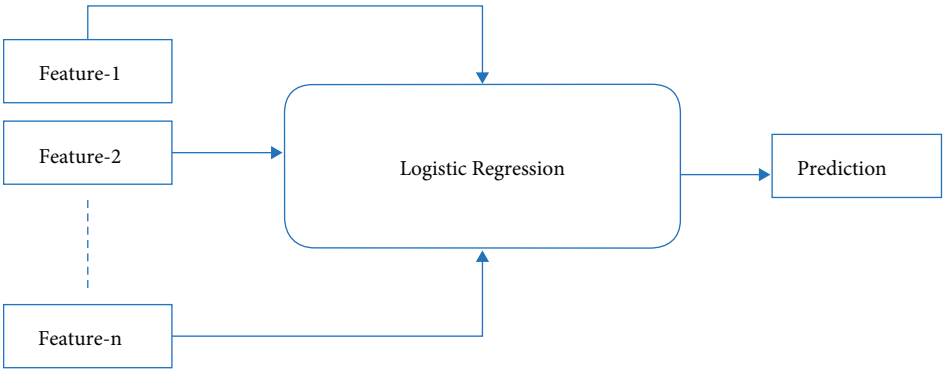


Figure 7: Structure of a logistic regression classifier.

shapes are all linked. The squares form the statures are proportional to the comparative class frequencies and recurrence densities for distinct classes. Figure 9 illustrates some important features of the histograms. A histogram depicts the dataset's proportions.

Figure 9 depicts the dataset's gender, age, hypertension, heart disease, ever married, average glucose level, and body mass index distributions. For the gender attribute, 0 means male and 1 means female. There are more female samples than male samples in this collection. However, based on the age distribution, it is obvious that the sample's average age is in the 40s, and the upper limit is approximately 60. When it comes to hypertension, 0 means the individual does not have it, while 1 means the person has it. The total number of individuals who are healthy and have no history of heart disease is achieved in this dataset. With regard to BMI and average glucose levels, Figure 10 shows the relationship between one feature and the target feature.

Figure 10 shows the relationship between gender and stroke, age and stroke, hypertension and stroke, heart disease and stroke, ever_married and stroke, avg_glucose_level and stroke, and BMI and stroke.

3.2. Visualization of Feature Selection. The process of feature selection is shown in Figure 11. Feature selection aids in comprehending how features are linked to one another.

Figure 11 shows that age, hypertension, avg_glucose_level, heart_disease, ever_married, and BMI are positively corelated with the target feature. However, gender is negatively corelated with stroke.

3.3. Evaluation of the Model

3.3.1. Random Forest (RF). Figure 12 depicts the classification report for the RF model.

In this case, the total F1-score obtained is 96 percent. The individual F1-scores for healthy people are 96 percent, while those who have had a brain stroke have 96 percent. This model achieved the highest accuracy after fine-tuning. Prior to fine-tuning, the model had an accuracy of 92 percent.

Figure 13 depicts the random forest model's prediction. The predicted outcome and the model's calculated performance are shown in the confusion matrix. There are 2707 accurate guesses and 113 erroneous predictions.

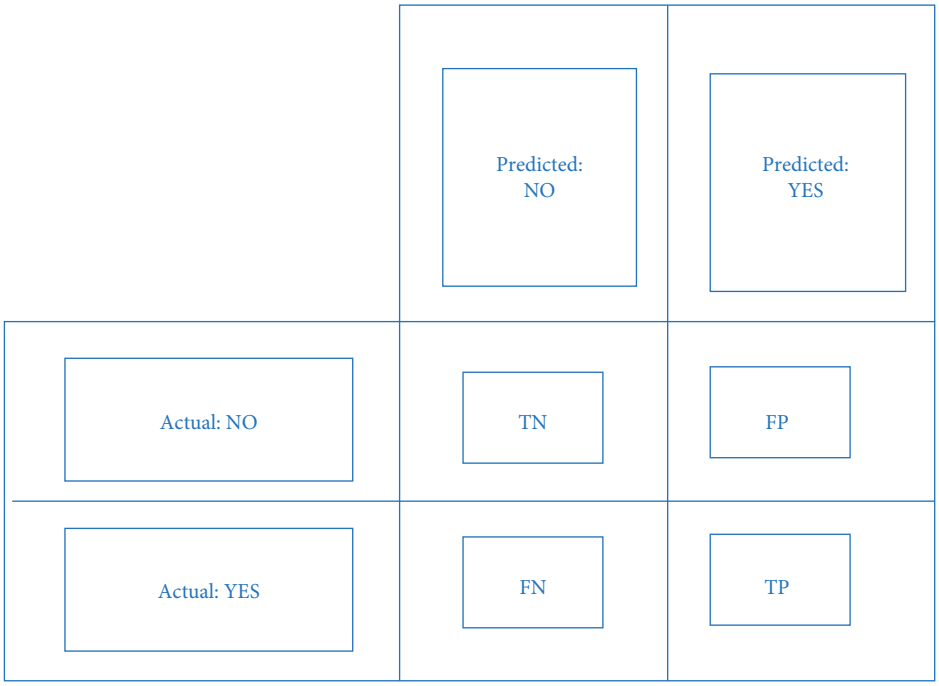


Figure 8: Block diagram of confusion matrix.

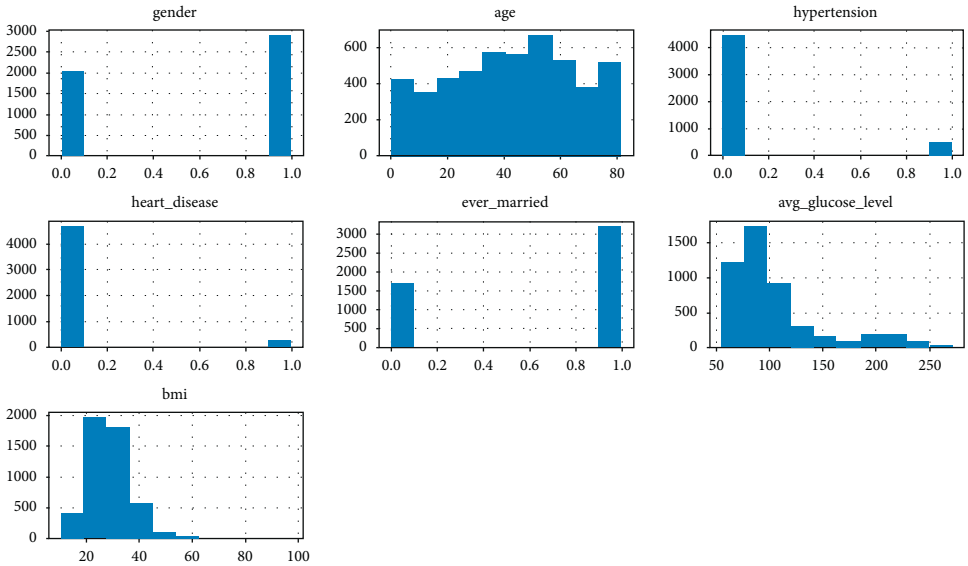


Figure 9: Histogram of some important features of the dataset.

3.3.2. Decision Tree. The classification report for the decision tree classification is shown in Figure 14.

The final F1-score in this case is 94 percent. An individual's F1-score is 94 percent for healthy individuals and 95 percent for those who have had a brain stroke. Also, the precision and recall are shown in Figure 14. A fine-tuned decision tree model has also been implemented. However, after fine-tuning, the accuracy did not improve.

Figure 15 depicts the DT model's prediction. There were 2664 accurate guesses and 156 erroneous predictions.

3.3.3. Voting Classifier. The classification report for the voting classifier is shown in Figure 16.

The total F1-score obtained in this case is 91 percent. The individual F1-scores are 91 percent for healthy people and 91 percent for those who have had a stroke. Also, the precision and recall are shown in Figure 16. Without any fine-tuning, this model achieved 91 percent accuracy.

The prediction made by the voting classifier is shown in Figure 17. The overall number of accurate guesses is 2565, while the total number of erroneous predictions is 255.

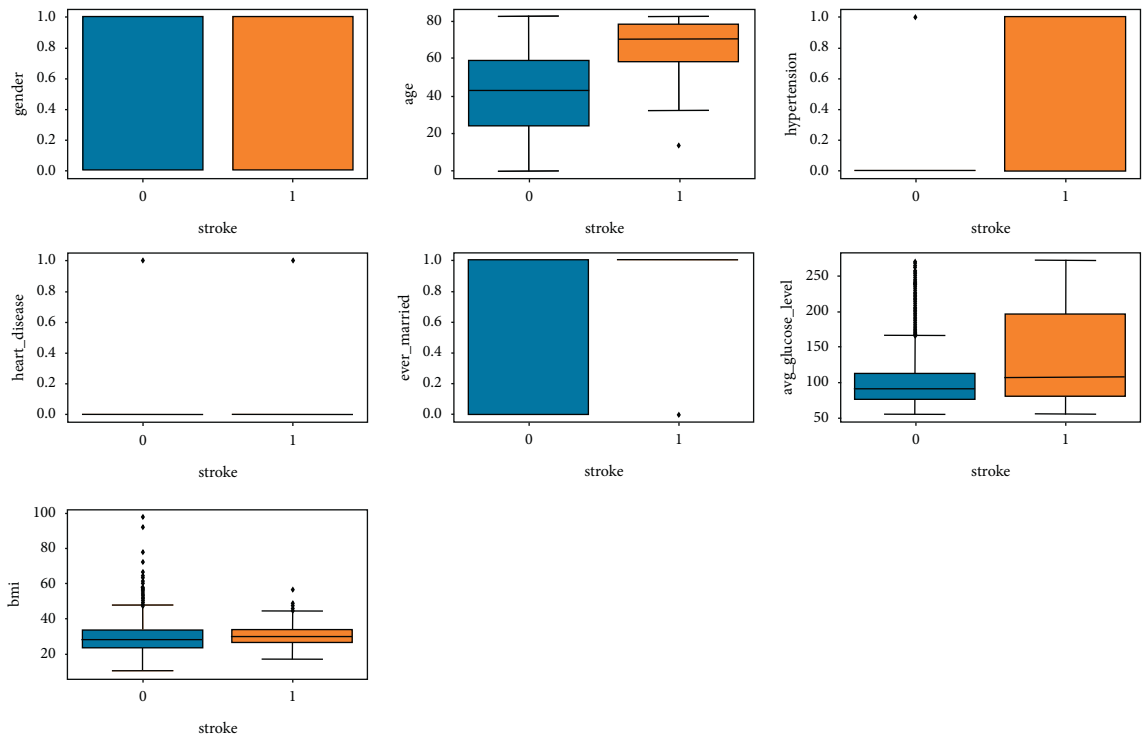


Figure 10: Relationship between some important features with the target feature.

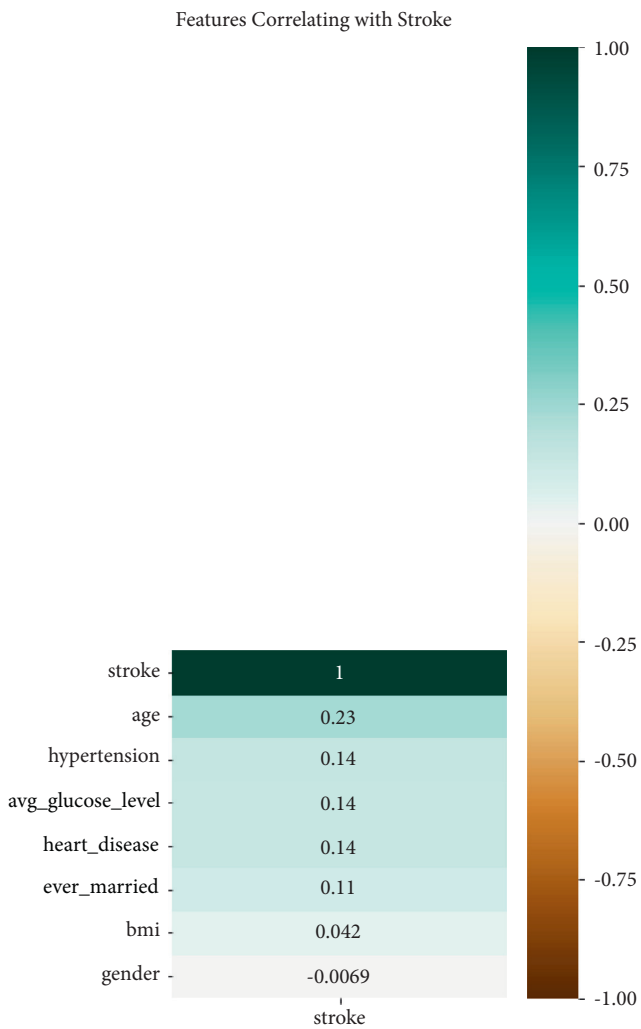


Figure 11: Features correlation with stroke.

```
[ ] print(classification_report(y_test,y_pred_rf))

              precision    recall  f1-score   support

           0       0.95      0.97      0.96      1407
           1       0.97      0.95      0.96      1413

    accuracy                           0.96      2820
   macro avg       0.96      0.96      0.96      2820
weighted avg       0.96      0.96      0.96      2820
```

Figure 12: Classification report of random forest.

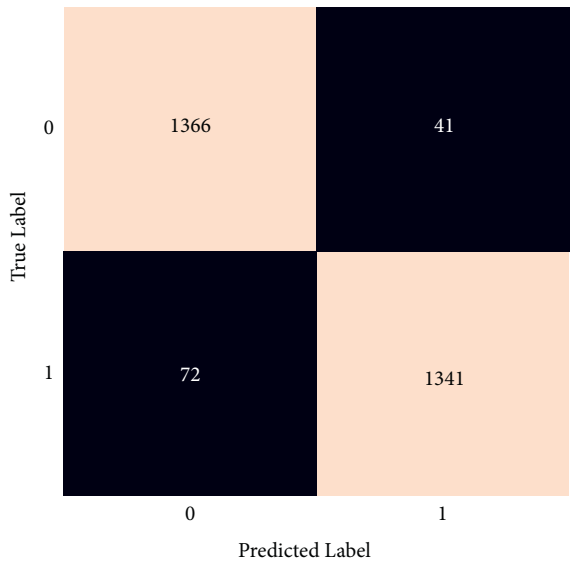


Figure 13: Confusion matrix of random forest.

```
[ ] y_pred_Clf = Clf.predict(X_test_s)
    print(classification_report(y_test, y_pred_Clf))

              precision    recall  f1-score   support

           0       0.95      0.94      0.94      1407
           1       0.94      0.95      0.95      1413

    accuracy                           0.94      2820
   macro avg       0.94      0.94      0.94      2820
weighted avg       0.94      0.94      0.94      2820
```

Figure 14: Classification report of decision tree.

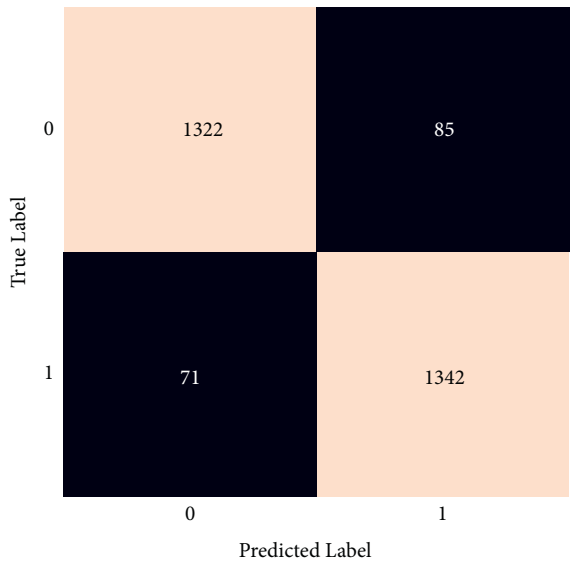


Figure 15: Confusion matrix of a decision tree.

```
y_pred_VC = VC.predict(X_test_s)
print(classification_report(y_test, y_pred_VC))
```

	precision	recall	f1-score	support
0	0.91	0.91	0.91	1407
1	0.91	0.91	0.91	1413
accuracy			0.91	2820
macro avg	0.91	0.91	0.91	2820
weighted avg	0.91	0.91	0.91	2820

Figure 16: Classification report of a voting classifier.

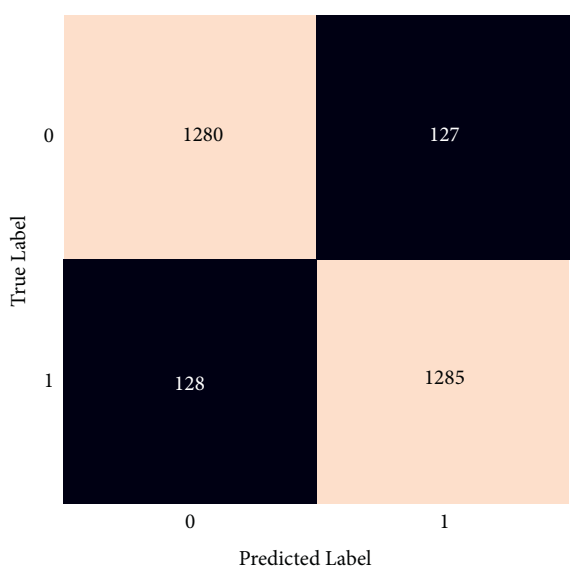


Figure 17: Confusion matrix of a voting classifier.

Table 1: Performance comparison.

This paper (model name)	Accuracy (%)	Reference paper (model name)	Accuracy (%)
Random forest	96	Ref [21] random forest	73
Decision tree	94	Ref [21] decision tree	77.6
Voting classifier	91	Ref [12] K-nearest neighbor	95
Logistic regression	79	Ref [21] logistic regression	77.6

3.4. Model Comparison. Table 1 shows a comparison of the models with those found in prior studies. The chart clearly demonstrates that of the various models included in the framework, the RF model is the most effective. In addition to having a higher F1-score, it has more precision and better recall and accuracy.

From Table 1, it is clear that all algorithms have an acceptable level of accuracy, but the random forest algorithm is a preferable option because of its higher level of accuracy. This paper achieved 96 percent accuracy using the RF algorithm, but in [21] the authors achieved only 73 percent accuracy. Also, using the decision tree algorithm, this paper achieved 94 percent accuracy, while the authors in [21] achieved 77.6 percent accuracy. Although the KNN algorithm has not been implemented in this research, ref [12] achieved 95 percent accuracy, which is higher than the voting classifier's accuracy (91 percent). However, in this paper, logistic regression performs poorly.

4. Conclusion

Stroke is a life-threatening medical illness that should be treated as soon as possible to avoid further complications. The development of an ML model could aid in the early detection of stroke and the subsequent mitigation of its severe consequences. The effectiveness of several ML algorithms in properly predicting stroke based on a number of physiological variables is investigated in this study. Random forest classification outperforms the other methods tested with a classification accuracy of 96 percent. According to the research, the random forest method outperforms other processes when cross-validation metrics are used in brain stroke forecasting. The future scope of this study is that using a larger dataset and machine learning models, such as AdaBoost, SVM, and Bagging, the framework models may be enhanced. This will enhance the dependability of the framework and the framework's presentation. In exchange for just providing some basic information, the machine learning architecture may help the general public in determining the likelihood of a stroke occurring in an adult patient. In an ideal world, it would help patients obtain early treatment for strokes and rebuild their lives after the event.

Data Availability

The data utilized to support this research findings are accessible online at https://www.kaggle.com/fedesoriano/stroke-prediction-dataset.

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

Acknowledgments

The authors are thankful for the support from Taif University Researchers Supporting Project (TURSP-2020/26), Taif University, Taif, Saudi Arabia.

References

[1] "Concept of stroke by healthline," [Online]. Available: https://www.cdc.gov/stroke/index.htm.

[2] "Statistics of stroke by Centers for disease control and prevention," [Online]. Available: https://www.cdc.gov/stroke/facts.htm.

[3] P. Govindarajan, R. K. Soundarapandian, A. H. Gandomi, R. Patan, P. Jayaraman, and R. Manikandan, "Classification of stroke disease using machine learning algorithms," *Neural Computing & Applications*, vol. 32, no. 3, pp. 817–828, 2020.

[4] L. Amini, R. Azarpazhouh, M. T. Farzadfar et al., "Prediction and control of stroke by data mining," *International Journal of Preventive Medicine*, vol. 4, no. 2, pp. S245–S249, 2013.

[5] S. M. Reza, M. M. Rahman, and S. A. Mamun, "A new approach for road networks-a vehicle xml device collaboration with big data," in *Proceedings of the International Conference On Electrical Engineering And Information And Communication Technology*, no. 1–5, Mirpur, Dhaka, April 2014.

[6] C. A. Cheng, Y. C. Lin, and H. W. Chiu, "Prediction of the prognosis of ischemic stroke patients after intravenous thrombolysis using artificial neural networks," *Studies in Health Technology and Informatics*, vol. 202, pp. 115–118, 2014.

[7] S. Cheon, J. Kim, and J. Lim, "The use of deep learning to predict stroke patient mortality," *International Journal of Environmental Research and Public Health*, vol. 16, no. 11, 2019.

[8] M. S. Zulfiker, N. Kabir, A. A. Biswas, P. Chakraborty, and M. M. Rahman, "Predicting students' performance of the private universities of Bangladesh using machine learning approaches," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 3, 2020.

[9] S. Rahman, T. Sharma, S. Reza, M. Rahman, M. Kaiser, and Nurjahan, "Pso-nf based vertical handoff decision for ubiquitous heterogeneous wireless network (uhwn)," in *Proceedings of the 2016 International Workshop On Computational Intelligence (IWCI)*, pp. 153–158, Dhaka, Bangladesh, December 2016.

[10] M. S. Singh and P. Choudhary, "Stroke prediction using artificial intelligence," in *Proceedings of the 2017 8th Annual Industrial Automation And Electromechanical Engineering Conference (IEMECON)*, pp. 158–161, Bangkok, Thailand, August 2017.

[11] C.-L. Chin, B.-J. Lin, G.-R. Wu et al., "An automated early ischemic stroke detection system using CNN deep learning algorithm," in *Proceedings of the 2017 IEEE 8th International Conference on Awareness Science and Technology (iCAST)*, pp. 368–372, Taichung, Taiwan, November 2017.

[12] S.-F. Sung, C.-Y. Hsieh, Y.-H. Kao Yang et al., "Developing a stroke severity index based on administrative data was feasible using data mining techniques," *Journal of Clinical Epidemiology*, vol. 68, no. 11, pp. 1292–1300, 2015.

[13] M. Monteiro, A. C. Fonseca, A. T. Freitas et al., "Using machine learning to improve the prediction of functional outcome in ischemic stroke patients," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 15, no. 6, pp. 1953–1959, 2018.

[14] T. Kansadub, S. Thammaboosadee, S. Kiattisin, and C. Jalayondeja, "Stroke risk prediction model based on demographic data," in *Proceedings of the 2015 8th Biomedical Engineering International Conference (BMEiCON)*, pp. 1–3, Pattaya, Thailand, November 2015.

[15] S. Y. Adam, A. Yousif, and M. B. Bashir, "Classification of ischemic stroke using machine learning algorithms," *International Journal of Computer Application*, vol. 149, no. 10, pp. 26–31, 2016.

[16] "Stroke prediction dataset," [Online]. Available: https://www.kaggle.com/fedesoriano/stroke-prediction-dataset.

[17] "Documentation for random forest classification from scikit-learn," org. [Online]. Available:https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html.

[18] "Documentation for decision tree classification from scikit-learn," org. [Online]. Available: https://scikit-learn.org/stable/modules/tree.html.

[19] Voting Classifier. [Online]. Available: https://towardsdatascience.com/custom-implementation-of-feature-importance-for-your-voting-classifier-model-859b573ce0e0.

[20] "Logistic regression in machine learning," [Online]. Available: .

[21] G. Sailasya and G. L. A. Kumari, "Analyzing the performance of stroke prediction using ML classification algorithms," *International Journal Of Advanced Computer Science And Applications*, vol. 12, no. 6, pp. 539–545, 2021.